

## 明細書

画像処理装置、画像処理方法、画像処理プログラム、及び記録媒体  
技術分野

【0001】 本発明は、単眼カメラによって撮像された三次元物体の動画像に  
5 画像処理を施す画像処理装置及び画像処理方法に関するものであり、特に、三次  
元物体の動画像から三次元物体の運動を推定する画像処理装置及び画像処理方法  
に関する。また、本発明は、コンピュータに当該画像処理方法を実行させる画像  
処理プログラム、及び当該画像処理プログラムが記録されたコンピュータ読取可  
能な記録媒体に関する。

### 背景技術

【0002】 三次元物体のローカルな動きを検出するための種々の画像処理方  
法がある。例えば、三次元物体である顔を対象として、カメラによって撮影され  
た顔の動画像から顔自体の姿勢や顔の各部の動き等の検出を行う画像処理方法が  
ある。このような画像処理方法には、眼の視線を求めるものがある（非特許文献  
15 1: J. Heinzmann and A. Zelinsky, "3-D fac  
ial pose and gaze point estimation u  
sing a robust real-time tracking par  
adigm", IEEE Int. Workshop on Automati  
c Face and Gesture Recognition, pp14  
20 2-147, 1998を参照）。非特許文献1には、単眼カメラで顔を撮像し、そ  
の動画像から顔を追跡しながら視線を求める方法が記載されている。非特許文献  
1に記載の方法によれば、顔の多数の特徴点がカルマンフィルタリングによって  
算出され、算出された特徴点を用いて顔の姿勢及び視線が求められる。

### 発明の開示

【0003】 上記した視線を求める方法では、顔姿勢や視線のパラメータが直  
接フィルタリングされるのではなく、特徴点にのみフィルタリングが行われるの

で、視線に対しての最適な推定を行えない。また、この方法では、眼の両端の特徴点と瞳の中心点との相対的な位置変化を用いて視線が推定される。したがって、特定の特徴点（すなわち、眼の両端の特徴点）に依存するので、視線の推定に限界がある。さらに、眼の両端の特徴点と瞳の中心点の相対的な位置変化を用いて  
5 視線が推定されるが、この方法は、カメラモデルのアフィン変換[Affine Transformation]を前提としている。そのため、この方法は、一般的な透視投影変換[Perspective Transformation]の下では成り立たないので、応用範囲が狭い。

【0004】 また、顔の場合、様々な表情変化があり、表情に応じて唇等の顔の各部が動く。そのため、顔の姿勢に対応しながら、顔の各部の運動を求めること  
10 とは非常に難しい。さらに、このような運動を行うのは、顔だけでなく、心臓等の様々な三次元物体に見られる。

【0005】 そこで、本発明は、三次元物体の全体的な運動に対応しつつ、三次元物体の各部のローカルな運動を高精度に求めることができる画像処理装置及び画像処理方法を提供することを目的とする。また、本発明は、コンピュータに  
15 当該画像処理方法を実行させる画像処理プログラム、及び当該画像処理プログラムを記録したコンピュータ読取可能な記録媒体を提供することを目的とする。

【0006】 本発明の一側面に係る画像処理装置は、単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を推定する画像処理装置である。この画像処理装置は、動画像のフレーム各々から、観測ベクトル抽出手段と、三次元モデル初期化手段と、運動推定手段とを備える。  
20

【0007】 観測ベクトル抽出手段は、所定の特徴点の画像面への投影座標を抽出する。三次元モデル初期化手段は、観測ベクトル抽出手段によって動画像の初期フレームから初期投影座標を抽出する。三次元モデル初期化手段は、第1のパラメータと第2のパラメータと初期投影座標とに基づいて前記所定の特徴点の  
25 モデル座標を算出するためのモデル座標演算式における該初期投影座標を観測ベクトル抽出手段によって動画像の初期フレームから抽出する。ここで、モデル座

標演算式は単眼カメラの逆投影に基づいており、第1のパラメータは所定の特徴点を含む部分のローカルな運動に依存しないパラメータであり、第2のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存するパラメータである。

【0008】 運動推定手段は、第1のパラメータ、第2のパラメータ、及び、  
5 第3のパラメータを含む状態変数の推定値を算出し、該状態変数の推定値に含まれる第2のパラメータに基づいて所定の特徴点の運動に関する出力値を出力する。ここで、第3のパラメータは、初期フレームとは異なる動画像の処理対象フレームが撮影された撮影時点での所定の特徴点の座標を三次元物体の運動モデルに基づきモデル座標から算出するための運動演算式におけるパラメータである。

10 【0009】 運動推定手段は、処理対象フレームが撮影された撮影時点における状態変数の予測値を状態遷移モデルに基づいて算出する。運動推定手段は、初期投影座標と、状態変数の予測値に含まれている第1のパラメータと第2のパラメータとを、モデル座標演算式に適用することによって、モデル座標の推定値を算出する。運動推定手段は、状態変数の予測値に含まれている第3のパラメータ  
15 とモデル座標の推定値とを運動演算式に適用することによって、撮影時点における所定の特徴点の座標の推定値を算出する。運動推定手段は、所定の特徴点の座標の推定値を、単眼カメラの観測モデルに基づく観測関数に適用することによって、所定の特徴点の観測ベクトルの推定値を算出する。運動推定手段は、処理対象フレームから所定の特徴点の投影座標を観測ベクトルとして観測ベクトル抽出  
20 手段によって抽出し、抽出した観測ベクトルと観測ベクトルの推定値とを用いて、状態変数の予測値をフィルタリングすることによって、撮像時点における前記状態変数の推定値を算出する。

【0010】 また、本発明の別の一側面に係る画像処理方法は、単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を  
25 推定する画像処理方法である。この画像処理方法は、三次元モデル初期化ステップと、運動推定ステップとを含む。

【0011】 三次元モデル初期化ステップでは、1のパラメータと第2のパラメータと初期投影座標とに基づいて所定の特徴点のモデル座標を算出するためのモデル座標演算式における該初期投影座標を動画像の初期フレームから抽出する。ここで、モデル座標演算式は単眼カメラの逆投影に基づいており、第1のパラメータは所定の特徴点を含む部分のローカルな運動に依存しないパラメータであり、第2のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存するパラメータである。

【0012】 運動推定ステップでは、第1のパラメータ、第2のパラメータ、及び、第3のパラメータを含む状態変数の推定値を算出し、該状態変数の推定値に含まれる第2のパラメータに基づいて所定の特徴点の運動に関する出力値を出力する。ここで、第3のパラメータは、初期フレームとは異なる動画像の処理対象フレームが撮影された撮影時点での所定の特徴点の座標を三次元物体の運動モデルに基づきモデル座標から算出するための運動演算式におけるパラメータである。

【0013】 運動推定ステップでは、処理対象フレームが撮影された撮影時点における状態変数の予測値を状態遷移モデルに基づいて算出する。運動推定手段は、初期投影座標と、状態変数の予測値に含まれている第1のパラメータと第2のパラメータとを、モデル座標演算式に適用することによって、モデル座標の推定値を算出する。運動推定ステップでは、状態変数の予測値に含まれている第3のパラメータとモデル座標の推定値とを運動演算式に適用することによって、撮影時点における所定の特徴点の座標の推定値を算出する。運動推定ステップでは、所定の特徴点の座標の推定値を、単眼カメラの観測モデルに基づく観測関数に適用することによって、所定の特徴点の観測ベクトルの推定値を算出する。運動推定ステップでは、処理対象フレームから所定の特徴点の投影座標を観測ベクトルとして抽出し、抽出した観測ベクトルと観測ベクトルの推定値とを用いて、状態変数の予測値をフィルタリングすることによって、撮像時点における前記状態変

数の推定値を算出する。

【0014】 また、本発明の他の一側面に係る画像処理プログラムは、上述した本発明の画像処理方法をコンピュータに実行させるためのプログラムである。また、本発明の更に別の一側面に係る記録媒体は、コンピュータ読取可能な記録媒体であって、上述した本発明の画像処理プログラムが記録されている。

【0015】 上述した本発明によれば、状態変数には、所定の特徴点のモデル座標を算出するためのパラメータが、ローカルな運動に依存しない第1のパラメータと、ローカルな運動に依存する第2のパラメータとに分離されている。したがって、特定の値に収束すべき第1のパラメータと、ローカルな運動に伴って変化する第2のパラメータとが分離されているので、正確にモデル座標を推定することができる。また、三次元物体の全体的な運動後の所定の特徴点の座標が、三次元物体の全体的な運動を算出するための第3のパラメータを用いて、このモデル座標から推定される。したがって、当該運動後の座標を、三次元物体の全体的な運動に対応しつつ、算出することができる。更に、推定した所定の特徴点の座標から算出される観測ベクトルの推定値と実測の観測ベクトルとを用いて、状態変数の予測値をフィルタリングすることによって、所定の特徴点の運動を推定するためのパラメータに直接フィルタリングを施すことができ、これらパラメータを含む状態変数の推定値を正確に推定することができる。故に、本発明によれば、第2のパラメータに基づいて所定の特徴点の運動に関する出力値を正確に算出することができる。

【0016】 上述した本発明において、第1のパラメータは、特定の値に収束する静的パラメータであり、第2のパラメータは、所定の特徴点を含む部分の運動によって変化する動的パラメータであることができる。静的パラメータは、前記画像面から前記所定の特徴点への奥行きであることが好適であり、動的パラメータは、前記所定の特徴点を含む部分の回転運動を特定するための回転パラメータであることが好適である。この場合に、回転パラメータは、前記所定の特徴点

を含む部分の中心を原点とする座標系における2つの座標軸に対して、該原点から該所定の特徴点へのベクトルが各々なす角度であることができる。

【0017】 上述した本発明において、第1のパラメータは、`rigid`パラメータであり、第2のパラメータは、`non-rigid`パラメータであることができる。`rigid`パラメータは、前記画像面から前記モデル座標までの奥行きであることが好適であり、`non-rigid`パラメータは、所定の特徴点を含む部分の運動による所定の特徴点の位置変化に関する変化量であることが好適である。

【0018】 また、上述した本発明において、運動モデル、三次元物体の回転及び並進運動に基づくモデルであってもよく、この場合には、第3のパラメータは、三次元物体の並進量を特定するための並進パラメータ及び該三次元物体の回転量を特定するための回転パラメータであることができる。

【0019】 また、上述した本発明において、フィルタリングとして拡張カルマンフィルタリングを適用することが好適である。

【0020】 また、本発明に係る画像処理装置は、単眼カメラで顔を撮像し、撮像した動画像から視線を求める画像処理装置であって、顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、静的パラメータ及び動的パラメータを推定することによって視線を求めることを特徴とする。

【0021】 この画像処理装置では、単眼カメラで顔を撮像した動画像に対して、顔画像上の瞳中心点における三次元構造を静的パラメータと動的パラメータによって定義する。さらに、画像処理装置では、フィルタリング等によって静的パラメータ及び動的パラメータを推定し、この推定値から視線方向を求める。このように、画像処理装置では、実際に起こりうる瞳の動きを静的パラメータや動的パラメータを用いてモデリングしているので、高精度に視線を推定することができる。また、この画像処理装置は、単眼カメラで、無制限の顔の動きに対応しながら視線の最適推定を行うことができる。

【0022】 本発明に係る画像処理方法は、単眼カメラで顔を撮像し、撮像した動画像から視線を求める画像処理方法であって、顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、静的パラメータ及び動的パラメータを推定することによって視線を求めることを特徴とする。

5 【0023】 なお、動画像は、時間的に連続した画像である。静的パラメータは、決まった値を元々有しているパラメータであり、例えば、眼球や瞳に関する値である。動的パラメータは、決まった値を有していないパラメータであり、例えば、顔や眼球の動きで変わる値である。

10 【0024】 また、本発明の画像処理プログラムは、この画像処理方法をコンピュータに実行させるプログラムである。また、本発明の記録媒体は、当該画像処理プログラムが記録されたコンピュータ読取可能な記録媒体である。

15 【0025】 本発明の上記画像処理装置、上記画像処理方法、画像処理プログラム、及び記録媒体では、静的パラメータは、カメラ座標系における瞳の奥行きであると好適である。さらに、本発明の上記画像処理装置及び上記画像処理方法では、動的パラメータは、眼球の回転パラメータであると好適であり、特に、眼球の回転パラメータは、眼球の座標系での2つの座標軸に対して回転可能である2自由度であると好適である。

20 【0026】 本発明に係る画像処理装置は、単眼カメラで三次元物体を撮像し、撮像した動画像から三次元物体の運動を求める画像処理装置であって、画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータで定義し、rigidパラメータ及びnon-rigidパラメータを推定することによって三次元物体の運動を求めることを特徴とする。

25 【0027】 この画像処理装置では、単眼カメラで三次元物体を撮像した動画像に対して、その画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータによって定義する。さらに、画像処理装置では、フィルタリング等によってrigidパラメータ及びnon-rigidパラメータ

タを推定し、この推定値から三次元物体の運動を求める。このように、画像処理装置では、実際に起こりうる三次元物体の特徴点の動きを *rigid* パラメータや *non-rigid* パラメータを用いてモデリングしているので、*non-rigid* な変化を伴う三次元物体の運動を高精度に推定することができる。また、この画像処理装置は、単眼カメラで、無制限の三次元物体の動きに対応しながら三次元物体の各部の運動の最適推定を行うことができる。例えば、この画像処理装置を顔の各部の運動を求めるために適用した場合、顔の並進や回転による動きに対応しながら、顔に表情変化がある場合の顔の各部（唇等）の運動を求めることができる。

【0028】 本発明に係る画像処理方法は、単眼カメラで三次元物体を撮像し、撮像した動画像から三次元物体の運動を求める画像処理方法であって、画像上の三次元物体の三次元構造を *rigid* パラメータと *non-rigid* パラメータで定義し、*rigid* パラメータ及び *non-rigid* パラメータを推定することによって三次元物体の運動を求めることを特徴とする。

【0029】 なお、三次元物体としては、様々な三次元物体に適用可能であり、特に、三次元物体全体で回転や並進の運動を行うとともに三次元物体の各部で *non-rigid* な変化をする三次元物体に適用可能であり、例えば、顔や心臓である。*rigid* パラメータは、三次元物体の三次元構造における *rigid* 構造を表すためのパラメータである。*non-rigid* パラメータは、三次元物体の三次元構造における *non-rigid* な変化を表すためのパラメータである。

【0030】 また、本発明の画像処理プログラムは、この画像処理方法をコンピュータに実行させるプログラムである。また、本発明の記録媒体は、当該画像処理プログラムが記録されたコンピュータ読取可能な記録媒体である。

【0031】 本発明の上記画像処理装置、上記画像処理方法、画像処理プログラム、及び記録媒体では、*rigid* パラメータは、モデル座標系における三次



元物体の特徴点の奥行きであると好適である。さらに、本発明の上記画像処理装置及び画像処理方法では、`non-rigid`パラメータは、モデル座標系における三次元物体の特徴点の変化量であると好適である。

【0032】 なお、上記した各画像処理方法では、上記した画像処理装置と同様の作用効果を奏する。

#### 図面の簡単な説明

図1は、本発明の実施の形態に係る画像処理システムの構成を示す図である。

図2は、カメラ座標系を示す図である。

図3は、顔が固定の場合のカメラ座標系、顔座標系、眼球座標系、視線座標系の関係を示す図である。

図4は、眼球モデルを示す図である。

図5は、`Dynamic Converging-Connected Gaze Model`を示す図である。

図6は、顔が動いた場合のカメラ座標系、顔座標系、眼球座標系、視線座標系の関係を示す図である。

図7は、本発明の実施の形態に係る画像処理方法の流れを示すフローチャートである。

図8は、三次元モデル初期化ステップの詳細な処理の流れを示すフローチャートである。

図9は、運動推定ステップの詳細な処理の流れを示すフローチャートである。

図10は、拡張カルマンフィルタを用いる処理の詳細な流れを示すフローチャートである。

図11は、本発明の実施の形態に係る画像処理プログラムの構成を、録媒体と共に示す図である。

図12は、記録媒体に記憶されたプログラムを実行するためのコンピュータのハードウェア構成を示す図である。

図 1 3 は、記録媒体に記憶されたプログラムを実行するためのコンピュータの斜視図である。

図 1 4 A は、図 1 の画像処理装置において視線を推定した場合の顔の画像例であり、顔姿勢が左方向で視線が左方向の場合の画像例である。

5 図 1 4 B は、図 1 の画像処理装置において視線を推定した場合の顔の画像例であり、顔姿勢が右方向で視線が左方向の場合の画像例である。

図 1 5 A は、顔の動き及び唇の動きが初期状態の場合における顔の唇の各特徴点を模式的に示す図である。

10 図 1 5 B は、唇のみが動いた場合における顔の唇の各特徴点を模式的に示す図である。

図 1 5 C は、図 1 5 B に示す唇の動きに顔の動きを伴った場合における各特徴点を模式的に示す図である。

図 1 6 は、モデル座標系を示す図である。

図 1 7 は、顔が並進及び回転運動した場合のモデル座標系を示す図である。

#### 15 発明を実施するための最良の形態

【0033】 以下、図面を参照して、本発明の好適な実施の形態を説明する。

【0034】 本実施の形態は、本発明を、単眼カメラで撮像された三次元物体としての顔の動画像に画像処理を施す画像処理システムの画像処理装置に適用したものである。本実施の形態に係る画像処理装置は、コンピュータ上に構成され、  
20 顔の動画像から顔における運動を推定することができる装置である。

【0035】 本実施の形態には、2つの形態がある。第1の実施の形態は、動画像の各フレームが撮影された時点における眼の視線方向を顔の各部のローカルな運動として推定する画像処理装置である。第2の実施の形態は、顔の各部（唇等）の運動を推定する画像処理装置である。なお、本実施の形態は、顔の動きに  
25 は制限がなく、第1の実施の形態では右眼の視線方向と左眼の視線方向とは同じ方向として視線の推定を行う。

【0036】 [第1の実施の形態]

【0037】 以下、本発明の第1の実施の形態について説明する。図1は、本発明の実施の形態に係る画像処理システムの構成を示す図である。

【0038】 図1に示す画像処理システム1は、透視投影変換のカメラモデルを用い、非接触で視線推定を行う。画像処理システム1では、画像取得装置2が画像処理装置3に接続されている。画像処理装置3は、画像取得装置2によって撮影された顔の動画像データを取得する。そして、画像処理システム1では、画像処理装置3によって、その取得した動画像データから視線の最適推定が行われる。画像処理システム1では、必要に応じて、画像処理の過程や結果を出力するためにモニタ（図示せず）等の出力装置が、画像処理装置3に接続されている。なお、画像処理システム1は、少なくとも視線の推定を行うが、必要に応じて、顔の姿勢の推定等の顔に対する他の画像処理も行う。

【0039】 画像取得装置2は、CCD[Charge Coupled Device]カメラ（図示せず）及びデータ変換装置（図示せず）からなる。CCDカメラは、単眼カメラであり、撮像素子としてCCDを用いている。CCDカメラでは、カメラのレンズから入射する光がCCDに入射し、CCDによって光の情報がデジタルデータに変換されて画像データとして出力される。また、CCDカメラは、撮像対象を時間的に連続して撮像し、時間的に連続した画像（動画像）データを出力する。データ変換装置は、CCDカメラから入力された動画像データの形式を画像処理装置3で取り扱うことができる形式に変換する。画像取得装置2は、CCDカメラによって被験者の顔全体を撮像し、データ変換装置によってその動画像データを所定の形式に変換して画像処理装置3に出力する。

【0040】 以下、図1～図6を参照して、画像処理装置3について説明する。図2は、カメラ座標系を示す図である。図3は、顔が固定の場合のカメラ座標系、顔座標系、眼球座標系、視線座標系の関係を示す図である。図4は、眼球モデルを示す図である。図5は、Dynamic Converging-Conne

cted Gaze Modelを示す図である。図6は、顔が動いた場合のカメラ座標系、顔座標系、眼球座標系、視線座標系の関係を示す図である。

【0041】 画像処理装置3は、パーソナルコンピュータ等のコンピュータに構成されている。画像処理装置3は、三次元モデル初期化部（三次元モデル初期化手段）10、推定部（運動推定手段）11、観測ベクトル抽出部（観測ベクトル抽出手段）12を備えている。三次元モデル初期化部10及び推定部11は、画像処理装置3用のアプリケーションプログラムをコンピュータ上で実行することによって構築される。

【0042】 なお、画像処理装置3は、 $N+2$ 個（ $i=1, 2, \dots, N+2$ ）の特徴点を取り扱うが、そのうちの $N$ 個（ $i=1, 2, \dots, N$ ）が瞳の中心点以外の顔の特徴点であり、2個（ $i=N+1, N+2$ ）が左右の瞳の中心点である。

【0043】 画像処理装置3では、図3及び図6に示す4つの座標系を定義している。図3～図6において、 $(X_c, Y_c, Z_c)$ は画像処理を行う上での基準の座標系であるカメラの座標系の互いに直交する三軸であり、 $(X_h, Y_h, Z_h)$ は顔の座標系の互いに直交する三軸であり、 $(X_e, Y_e, Z_e)$ は眼球の座標系の互いに直交する三軸であり、 $(X_g, Y_g, Z_g)$ は視線（瞳）の座標系の互いに直交する三軸である。

【0044】 観測ベクトル抽出部12は、顔の特徴点の画像面における投影座標（ $u, v$ ）を動画像の各フレームから抽出する。観測ベクトル抽出部12は、例えば、固有空間法と称される方法によって、顔の複数の特徴点各々の投影座標を抽出することができる。固有空間法の詳細については、M.A.Turk and A.P.Pentland, "Face recognition using eigenfaces," Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp.586-591, 1991."を参照されたい。以下、固有空間法に基づき顔の特徴点の投影座標を抽出する方法について簡単に説明する。

【0045】 固有空間法では、処理対象のフレームを全探索し、特徴点ごとに生成された複数の固有空間に、探索位置における部分画像を投影した後、逆投影することによって、複数の近似画像を求める。複数の固有空間は各々、複数の被験者を予め撮影した画像から、対応の特徴点の部分画像を用いて生成されている。

5 固有空間法では、全探索によって得られる全ての近似画像を用いてマッチングを行うことによって、最も誤差の少ない近似画像を特定し、特定した近似画像の基の部分画像の処理対象のフレームにおける位置が、抽出すべき特徴点の投影座標とされる。

10 【0046】 この固有空間法に基づく処理を実現するために、観測ベクトル抽出部12は、予め複数の被験者の顔の画像を取得し、各々が対応の特徴点の部分画像から形成される固有空間Wを生成している。

15 【0047】 ここで、特徴点の番号を  $i$  ( $i$  は 1 以上の整数)、被験者の番号を  $j$  ( $j$  は 1 以上の整数) とする。特徴点番号  $i$  の部分画像  $L_i$  の共分散行列  $Q_i$  は、 $j = 1 \sim n$  の部分画像  $L_{ij}$  から、次式 (a-1) 及び次式 (a-2) によって求められる。

$$Q_i = \frac{\sum_{j=1}^n (L_{ij} - \bar{L}_i)(L_{ij} - \bar{L}_i)}{n-1} \quad \dots(a-1)$$

$$\bar{L}_i = \frac{\sum_{j=1}^n L_{ij}}{n} \quad \dots(a-2)$$

20 【0048】 被験者の数  $n$  が  $L$  の次元数より小さいとすると、共分散行列  $Q_i$  の固有ベクトルは  $V_{ij}$ 、固有値は  $\lambda_{ij}$  ( $j = 1, \dots, n$ ) となる。ここで、 $\lambda_1 > \dots > \lambda_n$  とすると、近似した  $L_i$  の固有空間  $W_i$  は、次式 (a-3) によって定義される。

$$W_i = (v_1, \dots, v_d), \quad \frac{\sum_{j=1}^d \lambda_{ij}}{\sum_{j=1}^n \lambda_{ij}} \quad \dots (a-3)$$

【0049】 観測ベクトル抽出部12は、式(a-1)、(a-2)、及び(a-3)の演算によって複数の固有空間 $W_i$ を生成し、各々の固有空間 $W_i$ の固有ベクトル及び固有値を記憶している。

5 【0050】 観測ベクトル抽出部12は、処理対象のフレームを全探索し、探索位置における部分画像 $I$ を切り出し、該部分画像 $I$ を上記の固有空間 $W_i$ の全てに次式(a-4)の演算によって投影し、固有空間各々における点 $c_i$ を求める。

$$c_i = W_i^T (I - \overline{L_i}) \quad \dots (a-4)$$

10 【0051】 観測ベクトル抽出部12は、全ての投影点 $c_i$ を次式(a-5)の演算によって逆投影し、対応の投影点 $c_i$ に対する近似画像 $I^A_i$ を各々求める。

$$I^A_i = W_i c_i + L_i \quad \dots (a-5)$$

【0052】 観測ベクトル抽出部12は、全探索の結果得られる全ての近似画像の誤差 $E$ を次式(a-6)の演算によって求め、各特徴点に関して最小の誤差となる近似画像の基となった部分画像の座標を対応の特徴点の投影座標とする。

$$15 \quad E = \| I^A_i - I^B_i \|, \quad I^B_i = W^T (W (I^A_i - \overline{L_i})) + \overline{L_i} \quad \dots (a-6)$$

20 【0053】 観測ベクトル抽出部12は、動画像の初期フレームに対しては、固有空間法を用いて、全ての特徴点の投影座標を抽出する。観測ベクトル抽出部12は、その他のフレームに対しては、初期フレームにおける投影座標近傍の部分画像を用いたブロックマッチングによって、投影座標を抽出することができる。なお、観測ベクトル抽出部12は、全てのフレームに対して、固有空間法を用いて投影座標を抽出してもよい。

【0054】 以下、三次元モデル初期化部10について説明する。三次元モデル初期化部10では、初期フレームに基づいて、顔の特徴点と瞳の特徴点（中心

点)の三次元構造を定義する。初期フレームは、動画像における時間的に最初のフレームであり、画像処理を行う上での基準となるフレームである。初期フレームでは、顔が固定であり、顔座標系 ( $X_h, Y_h, Z_h$ ) がカメラ座標系 ( $X_c, Y_c, Z_c$ ) に一致している。

- 5 【0055】 三次元モデル初期化部10は、まず、観測ベクトル抽出部12を動作させて、初期フレームの画像からN個の顔の特徴点の画像面への投影座標 ( $u_i, v_i$ ) を抽出する。このN個の投影座標は、初期投影座標として、後述する処理に用いられる。

- 10 【0056】 三次元モデル初期化部10は、初期フレームの画像からN個の顔の特徴点の座標 ( $x_i, y_i, z_i$ ) ( $i = 1, 2, \dots, N$ ) を抽出する。三次元モデル初期化部10は、抽出した顔の特徴点の座標 ( $x_i, y_i, z_i$ ) から次式 (1) に示すモデル座標演算式によって顔の三次元構造を定義する。顔の特徴点 ( $x_i, y_i, z_i$ ) は、カメラ座標上の位置である。

$$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} u_i \\ v_i \\ 0 \end{pmatrix} + \alpha_i \begin{pmatrix} u_i \beta \\ v_i \beta \\ 1 \end{pmatrix} \quad \dots (1)$$

$i = 1, 2, \dots, N$

- 15 【0057】 式 (1) において、 $i$  は、顔のN個の特徴点の番号を表し、( $u_i, v_i$ ) は、顔の特徴点をCCD上に投影した初期投影座標であり、初期フレームで求めた後はその値に固定される。 $\alpha_i$  は、顔の特徴点の座標 ( $x_i, y_i, z_i$ ) からCCDまでの距離であり、特徴点の奥行きを示している。 $\beta$  は、CCDカメラの焦点距離  $f$  の逆数 ( $f^{-1}$ ) である。

- 20 【0058】 式 (1) によって定義される顔の特徴点の座標 ( $x_i, y_i, z_i$ ) は、モデル座標である。モデル座標は、初期フレームが撮影された時点での顔とカメラとの位置関係が固定されている状態での特徴点の座標であり、顔の三次元構造を定義する。すなわち、三次元モデル初期化部10は、初期投影座標 ( $u_i, v_i$ ) を抽出する。また、三次元モデル初期化部10は、式 (1) の  $\alpha_i$  の初期値

を設定することによって、顔の三次元構造（三次元モデル）の初期化を行う。なお、式（１）の $\alpha_i$ の初期値は、任意の値であることができ、例えば、経験に基づく値が設定される。

【００５９】 なお、図２には、カメラ座標系が示されており、 $(0, 0, -f)$ がＣＣＤカメラの焦点（カメラのレンズの中心）であり、焦点距離 $f$ 離れた位置に画像面（すなわち、レンズからの光を受けるＣＣＤ）が位置している。カメラ座標系の原点 $(0, 0, 0)$ は、画像面上にある。画像面に顔の特徴点 $(x_c, y_c, z_c)$ が投影されると、その投影座標は、 $(u, v, 0)$ となる。カメラ座標系の $Z$ 軸方向において、原点 $(0, 0, 0)$ から特徴点 $(x_c, y_c, z_c)$ までの奥行きが $\alpha$ となる。このカメラ座標系では、焦点距離 $f$ と奥行き $\alpha$ とが分離されている。

【００６０】 次に、三次元モデル初期化部１０では、初期フレームの画像から左右の瞳の中心点の画像面への投影座標 $(u_i, v_i)$  ( $i = N+1, N+2$ )を抽出し、これを瞳の中心点の初期投影座標とする。

【００６１】 三次元モデル初期化部１０は、抽出した瞳の中心点の初期投影座標 $(u_i, v_i)$ 、眼球の回転角、カメラ座標系上の奥行きによって、瞳の中心点の三次元構造を定義する。そのために、三次元モデル初期化部１０では、カメラ座標系 $(X_c, Y_c, Z_c)$ 、顔座標系 $(X_h, Y_h, Z_h)$ 、眼球座標系 $(X_e, Y_e, Z_e)$ 、視線座標系 $(X_g, Y_g, Z_g)$ が定義されている（図３参照）。なお、画像処理装置３では、初期フレームにおいて、眼球座標系 $(X_e, Y_e, Z_e)$ とカメラ座標系 $(X_c, Y_c, Z_c)$ 及び顔座標系 $(X_h, Y_h, Z_h)$ とは $Z$ 軸が平行で向きが逆方向と仮定している。

【００６２】 顔が固定で眼球が回転した場合、視線座標系 $(X_g, Y_g, Z_g)$ は、顔座標系 $(X_h, Y_h, Z_h)$ に対して次式（２）、（２－１）、（２－２）及び（２－３）のように表される。顔座標系 $(X_h, Y_h, Z_h)$ （カメラ座標系 $(X_c, Y_c, Z_c)$ ）から視線座標系 $(X_g, Y_g, Z_g)$ への変換 $T_h^g$ は、顔座標系 $(X_h,$



$Y_h, Z_h$ ) (カメラ座標系  $(X_c, Y_c, Z_c)$ ) から眼球座標系  $(X_e, Y_e, Z_e)$  への変換  $T_h^e$  と眼球座標系  $(X_e, Y_e, Z_e)$  から視線座標系  $(X_g, Y_g, Z_g)$  への変換  $T_e^g$  との積によって表される (図3参照)。

$$T_h^g = T_h^e T_e^g = \begin{pmatrix} R_{Xh, \pi} & t_h \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R_e(\phi, \theta) & 0 \\ 0 & 1 \end{pmatrix} \dots (2)$$

$$t_h = \begin{pmatrix} u_i(1 + \beta \alpha_i) \\ v_i(1 + \beta \alpha_i) \\ \alpha_i + r_{eye} \end{pmatrix} \dots (2-1)$$

$$R_{Xh, \pi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \pi & -\sin \pi \\ 0 & \sin \pi & \cos \pi \end{pmatrix} \dots (2-2)$$

$$R_e(\phi, \theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ -\sin \phi \sin \theta & \cos \phi & \sin \phi \cos \theta \\ -\cos \phi \sin \theta & -\sin \phi & \cos \phi \cos \theta \end{pmatrix} \dots (2-3)$$

【0063】 式(2)、(2-1)、(2-2)及び(2-3)において、 $i$ は、 $N+1$ 、又は $N+2$ であり、瞳の特徴点(中心点)の番号を示している。 $T_h^g$ は、視線座標系の顔座標系に対する変換( $4 \times 4$ の行列)である。 $T_h^e$ は、眼球座標系の顔座標系に対する変換( $4 \times 4$ の行列)である。変換 $T_h^e$ は、式(2-1)による眼球座標系の顔座標系に対する並進移動 $t_h$ ( $3 \times 1$ の並進行列)と式(2-2)による眼球座標系の顔座標系に対する回転 $R_{Xh, \pi}$ ( $3 \times 3$ の回転行列)によって表される(図3参照)。並進移動 $t_h$ は、顔座標系 $(X_h, Y_h, Z_h)$ (カメラ座標系 $(X_c, Y_c, Z_c)$ )から眼球座標系 $(X_e, Y_e, Z_e)$ への原点の移動である。回転 $R_{Xh, \pi}$ は、 $Z$ 軸に対する $180^\circ$ 回転である。式(2-1)において、 $u_i$ 及び $v_i$ は、瞳の中心点の画像面(CCD)への投影座標値であり、 $r_{eye}$ は眼球の半径であり、 $\alpha_i$ は、画像面(CCD)から瞳の中心点までの距離(瞳の奥行き)である。

【0064】  $T_e^g$ は、視線座標系の眼球座標系に対する変換( $4 \times 4$ の行列)

であり、式(2-3)による視線座標系の眼球座標系に対する回転 $R_e$  ( $3 \times 3$ の回転行列)によって表される(図3参照)。式(2-3)において、 $\phi$ 及び $\theta$ は、眼球座標系における瞳の中心点の動き(眼球の回転)を示す回転角度である。回転 $R_e$ は、図4に示すように、眼球の回転による瞳の中心点の $I_0$ から $I$ への回転である。この瞳の中心点である $I$ は、式(3)によって表される。

$$\begin{pmatrix} I \\ 1 \end{pmatrix} = T_h^g \begin{pmatrix} I_0 \\ 1 \end{pmatrix} = T_h^g \begin{pmatrix} 0 \\ 0 \\ r_{eye} \\ 1 \end{pmatrix} \quad \dots (3)$$

【0065】 式(3)において、 $I_0$ は、初期フレームでの眼球座標系における瞳の中心点であり、 $I$ は、顔座標系からの瞳の中心点である。 $I_0$ は、図4に示すように、眼球座標系( $X_e, Y_e, Z_e$ )において $Z$ 軸上にあり、 $Z$ 軸上における眼球の半径 $r_{eye}$ に位置する。眼球の回転による瞳の中心点の $I_0$ から $I$ への移動は、図4に示すように、2つの回転角 $\theta, \phi$ によって規定される。つまり、眼球座標系( $X_e, Y_e, Z_e$ )のうちの $Y_e$ 軸と $X_e$ 軸に対して回転する角度によって規定されている。この2つの軸に対して回転可能となっている回転角を動的パラメータとする。 $\theta, \phi$ は、 $Y_e$ 軸と $X_e$ 軸に対して回転可能であることから2自由度である。

【0066】 図4に示す $I_0$ を中心とした $r_{iris}$ は、瞳の半径である。瞳の半径 $r_{iris}$ を画像面(CCD)上に投影すると、式(4-1)に示すようになる。さらに、眼球の半径 $r_{eye}$ を画像面上に投影した瞳の半径 $s_i$ によって表すと、次式(4-2)に示すようになる。

$$s_i = \frac{r_{iris}}{1 + \beta \alpha_i} \quad \dots (4-1)$$

$$r_{eye} = \frac{r_{eye}}{r_{iris}} \times \frac{r_{iris}}{s_i} \times s_i = \lambda s_i (1 + \beta \alpha_i) \quad \dots (4-2)$$

【0067】 式(4-1)及び式(4-2)において、 $i$ は $N+1$ 又は $N+2$

であり、 $s_i$ は、画像面（CCD）上に投影した瞳の半径である。式（4-2）における $\lambda$ は、眼球と瞳との半径比であり、一般的に人によって差が少ないので一定値と仮定することができる。さらに、左右の眼球の半径も同じ値と仮定すると、次式（5）が成立する。

$$s_{N+1}(1 + \beta \alpha_{N+1}) = s_{N+2}(1 + \beta \alpha_{N+2}) \quad \dots (5)$$

【0068】 式（5）において、下付き文字の $N+1$ 、 $N+2$ のうちの一方が左の眼球（瞳）を示し、他方が右の眼球（瞳）を示す。

【0069】 三次元モデル初期化部10では、式（2）～式（5）から導かれる次式（6）に示すモデル座標演算式によって、顔が固定の場合の眼球運動による瞳の中心点の三次元構造を定義する。

$$I = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} (u_i + \lambda s_i \sin \theta) (1 + \beta \alpha_i) \\ (v_i + \lambda s_i \sin \phi \cos \theta) (1 + \beta \alpha_i) \\ (\alpha_i + \lambda s_i (1 - \cos \phi \cos \theta) (1 + \beta \alpha_i)) \end{pmatrix} \quad \dots (6)$$

【0070】 式（6）において、 $i$ は、 $N+1$ 又は $N+2$ であり、 $x_i$ 、 $y_i$ 、 $z_i$ は、顔が固定の場合のカメラ座標系上の瞳の中心点の座標値である。すなわち、 $I$ は、瞳の中心点のモデル座標であり、 $x_i$ 、 $y_i$ 、 $z_i$ は瞳の中心点のモデル座標の座標値である。また、 $\alpha_i$ は、次式（6-1）に示す関係をもつ。

$$\alpha_{N+2} = \frac{1}{\beta} \left( \frac{s_{N+1}}{s_{N+2}} - 1 \right) + \frac{s_{N+1}}{s_{N+2}} \alpha_{N+1} \quad \dots (6-1)$$

【0071】 式（6）において、瞳中心点の初期投影座標（ $u_i$ 、 $v_i$ ）は、初期フレームから求められた固定値である。半径比 $\lambda$ は、実際には人によって異なるがその差が少ないので、人体工学によって求められる一定値を用いる。ちなみに、半径比 $\lambda$ は、実際に測定した値を用いてもよい。回転角（ $\theta$ 、 $\phi$ ）は、眼球の回転運動によって変わり、未知である。回転角（ $\theta$ 、 $\phi$ ）は、値が決まっていないので、動的パラメータとする。焦点距離 $\beta$ は、CCDカメラによって変わり、未

知である。瞳中心点の奥行き $\alpha_1$ は、決まった値を元々有しているが、未知である。奥行き $\alpha_1$ は、ある決まった値に収束するので、静的パラメータとする。

【0072】 このように、三次元モデル初期化部10は、瞳の中心点の初期投影座標( $u_1, v_1$ )を抽出し、静的パラメータ $\alpha_1$ 、動的パラメータ $\phi$ 及び $\theta$ の初期値を設定することによって、顔の中心点の三次元構造(三次元モデル)の初期化を行う。式(6)の $\alpha_1$ 、 $\phi$ 、 $\theta$ の初期値は、任意の値であることができ、例えば、経験に基づく値が設定される。

【0073】 なお、瞳中心点の奥行き $\alpha_1$ は、式(5)から、式(6-1)に示すように、左右の瞳の一方の奥行き $\alpha_{N+2}$ を他方の奥行き $\alpha_{N+1}$ で表すことができる。したがって、他方の奥行き $\alpha_{N+1}$ を推定すれば、一方の奥行き $\alpha_{N+2}$ を求めることができる。

【0074】 以下、推定部11について説明する。推定部11では、顔の運動モデルと眼球の運動モデルが定義されており、この2つの運動モデルからDynamic Converging-Connected Gaze Model(以下、DCCGMと記載する)が生成される。推定部11は、DCCGMに対して拡張カルマンフィルタを適用し、視線を推定する。

【0075】 具体的に、推定部11は、上記の静的パラメータ $\alpha$ 及び動的パラメータ $\theta$ 、 $\phi$ を含む状態変数の予測値を算出し、状態変数の予測値に含まれる静的パラメータ $\alpha$ 及び動的パラメータ $\theta$ 、 $\phi$ 、並びに初期投影座標( $u, v$ )を用いて推定されるモデル座標から運動モデルに基づき運動後の特徴点の座標の推定値を算出する。推定部11は、観測モデルに基づき、特徴点の座標の推定値を画面上に投影した投影座標、すなわち観測ベクトルの推定値を算出し、観測ベクトルの推定値と処理対象のフレームから抽出される観測ベクトルとを用いて、状態変数の予測値をフィルタリングすることによって、状態変数の推定値を算出する。このフィルタリングには、拡張カルマンフィルタが用いられる。推定部11は、状態変数の推定値に含まれている動的パラメータに基づいて、視線方向を抽出す

る。

【0076】 以下、より詳細に推定部11について説明する。推定部11では、顔の運動モデルとして、回転運動と並進運動とが想定されている。顔の動きをカメラ座標系 ( $X_c, Y_c, Z_c$ ) に対する回転運動と並進移動とで表すと、顔の特徴点の座標 ( $x_{ci}, y_{ci}, \beta z_{ci}$ ) は、次式 (7-1) の運動演算式によって表される。

$$\begin{pmatrix} x_{ci} \\ y_{ci} \\ \beta z_{ci} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} + R_c \bar{m}_i \quad \dots (7-1)$$

【0077】 式 (7-1) において、 $x_{ci}, y_{ci}, z_{ci}$  は各々、顔が動いた場合の顔の特徴点をカメラ座標系上に投影した座標の座標値であり、( $t_x, t_y, t_z$ ) はカメラ座標系 ( $X_c, Y_c, Z_c$ ) に対する顔座標系の並進移動量であり、 $3 \times 1$  の並進行列である。 $R_c$  はカメラ座標系 ( $X_c, Y_c, Z_c$ ) に対する顔座標系の回転を表す回転行列であり、 $3 \times 3$  の回転行列である。 $\bar{m}_i$  ベクトルは、顔の各特徴点のモデル座標のベクトル表現であり、 $i=1, \dots, N$  の場合 (瞳の中心点を除く場合) には、式 (1) を変形した次式 (7-1-1) によって表すことができる。なお、瞳の中心点のモデル座標は、式 (6) から求めることができる。

$$\bar{m}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} u_i + \alpha_i u_i \beta \\ v_i + \alpha_i v_i \beta \\ \alpha_i \end{pmatrix} \quad \dots (7-1-1)$$

【0078】 式 (7-1) における  $R_c$  は、次式 (7-1-2) から求めることができる。

$$R_c = R(\delta q) \cdot R_{old} \quad \dots (7-1-2)$$

式 (7-1-2) において、 $R_{old}$  は、処理対象のフレームに関連する時点である現時点より前の時点に対して求められた  $R_c$  であり、初期には、単位行列とする。

ることができる。R (q) は、次式 (7-1-3) によって、 $\delta q$  は、次式 (7-1-4) 及び (7-1-5) によって表すことができる。なお、(7-1-5) における  $w_x$ 、 $w_y$ 、 $w_z$  については後述する。

$$R(q) = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad \dots(7-1-3)$$

$$\delta q = \left( \sqrt{1-\varepsilon}, \frac{w_x}{2}, \frac{w_y}{2}, \frac{w_z}{2} \right)^T \quad \dots(7-1-4)$$

$$\varepsilon = \sqrt{\left( \frac{w_x}{2} \right)^2 + \left( \frac{w_y}{2} \right)^2 + \left( \frac{w_z}{2} \right)^2} \quad \dots(7-1-5)$$

【0079】 推定部11では、観測モデルとして画像面上への透視投影変換が想定されている。この顔の特徴点 ( $x_{ci}$ ,  $y_{ci}$ ,  $\beta z_{ci}$ ) を画像面 (CCD) 上に投影した投影座標の推定値、すなわち、 $o_i$  ベクトル (観測ベクトル) の推定値は、式 (7-1) から算出される顔の特徴点の座標値を用いる次式 (7-2) に示す観測関数  $f$  によって表すことができる。

$$\vec{o}_i = f(x_{ci}, y_{ci}, z_{ci}) = \begin{pmatrix} \frac{x_{ci}}{1 + \beta z_{ci}} \\ \frac{y_{ci}}{1 + \beta z_{ci}} \end{pmatrix} \quad \dots(7-2)$$

【0080】 ここで、全ての顔の特徴点 ( $i = 1, 2, \dots, N$ ) に対する観測ベクトルである  $O$  ベクトルは、式 (7-3) で表すことができる。このように、推定部11は、顔の特徴点に対する観測モデルを表す。

$$\bar{O} = \begin{pmatrix} \bar{o}_1 \\ \bar{o}_2 \\ \vdots \\ \bar{o}_N \end{pmatrix} \quad \dots (7-3)$$

【0081】 推定部11では、この顔の運動モデルの状態遷移表現（状態遷移モデル）における状態変数であるpベクトルが次式（8-1）で定義されている。pベクトルは、顔の姿勢と構造パラメータで構成される。推定部11では、顔の運動モデルの状態空間表現が次式（8-2）で定義されている。

$$\bar{p} = \begin{pmatrix} t_x \\ t_y \\ \beta t_z \\ w_x \\ w_y \\ w_z \\ \beta \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{N+1} \end{pmatrix} \quad \dots (8-1)$$

$$\bar{p}_t = A^{(1)} \bar{p}_{t-1} + \mu_t^{(1)} \quad \dots (8-2)$$

【0082】 式（8-1）において、 $(t_x, t_y, \beta t_z)$  は顔座標系  $(X_h, Y_h, Z_h)$  のカメラ座標系  $(X_o, Y_o, Z_o)$  に対する並進移動量であり、 $(w_x, w_y, w_z)$  は前フレームからのカメラ座標系  $(X_o, Y_o, Z_o)$  に対する各軸の回転量であり、回転行列  $R_o$  に関する量である。 $\beta$  はCCDカメラの焦点距離の逆数であり、 $\alpha_1 \sim \alpha_{N+1}$  は顔の特徴点と瞳中心点の各奥行きである。式（8-2）において、下付き文字のtは現時点（処理対象のフレームである現フレームに関連する時点）を示し、下付き文字のt-1は一時刻前（現フレームの前フレ

ームに関連する時点)を示す。 $A^{(1)}$ は現状態と一時刻前の状態との顔の運動特性を表す状態遷移行列であり、 $\mu_t^{(1)}$ は現状態と一時刻前の状態とのノイズを表す。 $\mu_t^{(1)}$ は、ゼロの平均とノイズの分散共分散行列 $Q^{(1)}$ で構成されるガウシアン分布のランダムベクトルである。

- 5 【0083】 推定部11では、眼球の運動モデルの状態遷移表現(状態遷移モデル)における状態変数である $q$ ベクトルが式(9-1)で定義されている。 $q$ ベクトルは、2自由度の回転パラメータで構成される。さらに、推定部11では、眼球の運動モデルの状態空間表現が式(9-2)で定義されている。

$$\bar{q} = \begin{pmatrix} \phi \\ \theta \end{pmatrix} \quad \dots \quad (9-1)$$

$$\bar{q}_t = A^{(2)} \bar{q}_{t-1} + \mu_t^{(2)} \quad \dots \quad (9-2)$$

- 10 【0084】 式(9-1)において、 $(\theta, \phi)$ は眼球の回転を表す2つの回転角である(図4参照)。式(9-2)において、 $q_t$ ベクトルは、現時点での眼球の運動モデルの状態変数であり、 $q_{t-1}$ ベクトルは、一時刻前の眼球の運動モデルの状態変数である。 $A^{(2)}$ は現状態と一時刻前の状態との眼球の運動特性を表す状態遷移行列であり、 $\mu_t^{(2)}$ は現状態と一時刻前の状態とのノイズを表す。 $\mu_t^{(2)}$ は、ゼロの平均とノイズの分散共分散行列 $Q^{(2)}$ で構成されるガウシアン分布のランダムベクトルである。

- 15 【0085】 式(1)と式(6)から判るように、顔の特徴点の三次元構造と瞳の中心点の三次元構造とは、初期フレームの画像座標(初期投影座標)を用いることによって定義される面では同じである。しかしながら、瞳の中心点の三次元構造には、それに動的な画像上の移動量 $(\theta, \phi)$ が加えられている。そこで、この眼球座標系 $(X_e, Y_e, Z_e)$ の運動に対応するために、推定部11は、図5に示すように、顔の運動モデルと眼球の運動モデルとを結合し、一点に集中させるDCCGMを生成する。図5に示す $O_t$ ベクトル( $t=1, 2, \dots$ )は、顔の運動モデルの状態変数 $p$ ベクトルと眼球の運動モデルの状態変数 $q$ ベクトル



を結合した観測ベクトルである。 $O_t$ ベクトルは、顔の特徴点 ( $i = 1, 2, \dots, N$ ) 及び瞳の中心点 ( $i = N+1, N+2$ ) の全てのモデル点を画像面 (CCD) 上に投影した座標 (投影座標) で構成される。

【0086】 推定部11では、式(8-1)、式(8-2)で示す顔の運動モデルと式(9-1)、式(9-2)で示す眼球の運動モデルとから、DCCGMの状態空間表現が次式(10-1)で定義されている。

$$\bar{X}_t = \begin{pmatrix} \bar{p}_t \\ \bar{q}_t \end{pmatrix} = \begin{pmatrix} A^{(1)} & 0 \\ 0 & A^{(2)} \end{pmatrix} \bar{X}_{t-1} + \begin{pmatrix} \mu_t^{(1)} \\ \mu_t^{(2)} \end{pmatrix} \quad \dots (10-1)$$

式(10-1)において、 $X_t$ ベクトルは、現時点での顔の運動モデルと眼球の運動モデルの状態変数を示す $p_t$ ベクトルと $q_t$ ベクトルからなるDCCGMの状態変数である。 $X_{t-1}$ ベクトルは、一時刻前のDCCGMの状態変数である。推定部11は、式(10-1)に基づく状態遷移方程式に基づき、一時刻前の状態変数の推定値 $X_{t-1}$ から現時点の状態変数の予測値 $X_t^-$ を算出する。

【0087】 さらに、推定部11では、観測ベクトルである $O_t$ ベクトルが次式(10-2)で定義されている。

$$\bar{O}_t = \begin{pmatrix} \bar{o}_1 \\ \bar{o}_2 \\ \vdots \\ \bar{o}_{N+1} \\ \bar{o}_{N+2} \end{pmatrix} = f(\bar{X}_t) + v_t \quad \dots (10-2)$$

式(10-2)において、 $f()$ は観測関数であり、 $v_t$ はノイズを表す。 $O_t$ ベクトルは、 $(N+2)$ 個分の $o_i$ ベクトル、すなわち $N$ 個の顔の特徴点及び2個の瞳の特徴点それぞれの観測ベクトルを含んでいる。

【0088】 推定部11では、式(1)で定義した顔の特徴点 ( $i = 1, 2, \dots, N$ ) のモデル座標と式(6)で定義した瞳の中心点 ( $i = N+1, N+2$ ) のモ

デル座標との剛性変換（回転と並進移動）を行う運動演算式である次式（11）によって、顔が動いた場合の三次元構造が定義されている。

$$\begin{pmatrix} x_{ci} \\ y_{ci} \\ z_{ci} \\ 1 \end{pmatrix} = T_c^h \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = \begin{pmatrix} R_c & t_c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} \dots (11)$$

【0089】 式（11）において、 $(x_i, y_i, z_i)$  は、モデル座標であり、  
 5  $(x_{ci}, y_{ci}, z_{ci})$  は、顔が動いた場合の顔の特徴点及び瞳の中心点をカメラ座標系上に投影した座標であり、 $T_c^h$  はカメラ座標系  $(X_c, Y_c, Z_c)$  から顔座標系  $(X_h, Y_h, Z_h)$  への変換（ $4 \times 4$  の行列）である。 $R_c$  はカメラ座標系に対する顔の回転を示す  $3 \times 3$  の回転行列であり、 $t_c$  はカメラ座標系に対する顔の位置移動を示す  $3 \times 1$  の並進行列である（図6参照）。この顔の動きには、  
 10 制限がなく、回転行列  $R_c$  と並進行列  $t_c$  によってあらゆる動きを表すことができる。なお、 $R_c$  及び  $t_c$  については、式（7-1-2）の  $R_c$ 、式（7-1）の  $(t_x, t_y, t_z)$  を参照されたい。

【0090】 推定部11は、式（11）を用いて、現時点での特徴点の座標の推定値を算出する。推定部11は、特徴点の座標の推定値を算出する際に、顔の  
 15 特徴点（ $i = 1, 2, \dots, N$ ）に関しては、状態変数の予測値  $X_t^-$  に含まれているパラメータと初期投影座標とを式（1）で示すモデル座標演算式に適用することによって得られるモデル座標  $(x_i, y_i, z_i)$  を式（11）に適用する。  
 瞳の中心点（ $i = N+1, N+2$ ）に関しては、推定部11は、状態変数の予測値  $X_t^-$  に含まれているパラメータと初期投影座標とを式（6）で示すモデル座標  
 20 演算式に適用することによって得られるモデル座標  $(x_i, y_i, z_i)$  を式（11）に適用する。

【0091】 推定部11では、式（11）で示す顔の特徴点及び瞳の中心点の座標  $(x_{ci}, y_{ci}, z_{ci})$  を画像面（CCD）上に投影した  $o_i$  ベクトル（観測ベクトル）が次式（12）によって定義されている。すなわち、推定部11は、

式(11)によって得られる顔の特徴点及び瞳の中心点の座標の推定値( $x_{ci}$ ,  $y_{ci}$ ,  $z_{ci}$ )を、式(12)の観測関数 $f$ に適用することによって、観測ベクトルの推定値を算出する。

$$\vec{o}_i = f(x_{ci}, y_{ci}, z_{ci}) = \begin{pmatrix} \frac{x_{ci}}{1 + \beta z_{ci}} \\ \frac{y_{ci}}{1 + \beta z_{ci}} \end{pmatrix} \quad \dots(12)$$

5 【0092】 このように、推定部11は、式(10-2)で示す観測ベクトル $O_t$ の推定値から状態変数 $X_t$ を推定するために、式(11)を用いて、顔が動いた場合の三次元構造、すなわち特徴点の座標の推定値を算出し、当該推定値を式(12)に適用することによって観測ベクトル $o_i$ の推定値を算出する。

10 【0093】 推定部11は、式(10-2)の観測ベクトル $O_t$ に対して拡張カルマンフィルタリングを施し、式(10-1)の状態変数である $X_t$ ベクトル(すなわち、顔の運動モデルの状態変数である $p$ ベクトル及び眼球の運動モデルである $q$ ベクトル)の各パラメータ( $\alpha_1 \sim \alpha_{N+2}$ ,  $\beta$ ,  $\theta$ ,  $\phi$ 等)を推定する。この観測過程では、式(12)によって得られる観測ベクトル $o_i$ の推定値及び式(11)によって得られる三次元構造が用いられる。

15 【0094】 より具体的に、推定部11は、拡張カルマンフィルタを用いた推定を実行する前に、次式(b-1)及び(b-2)に示す予測を実行する。

$$X_t^- = AX_{t-1} \quad \dots(b-1)$$

$$P_t^- = P_{t-1} + Q \quad (b-2)$$

20 式(b-1)において、 $A$ は式(10-1)における状態遷移行列であり、 $X_{t-1}$ は一時刻前(時刻 $t-1$ )の状態変数の推定値であり、 $X_t^-$ は、現時点(時刻 $t$ )の状態変数の予測値である。本実施の形態では、動画像のフレーム間のタイムインターバルが短く、一時刻前と現時点とでは顔の動きが無視できるものと仮定し、 $A$ として単位行列が用いられている。

【0095】 また、式（b-2）において、 $P_{t-1}$ は一時刻前（時刻 $t-1$ ）の推定誤差共分散行列の推定値であり、 $P_t^-$ は現時点（時刻 $t$ ）の推定誤差共分散行列の予測値であり、 $Q$ は式（10-1）の状態遷移モデルにおけるノイズの分散共分散行列である。ここで、 $P_{t-1}$ の要素の初期値は、任意の値であることができ、例えば、0に設定される。ノイズの分散共分散行列 $Q$ は、フレーム間での顔の動きから経験的に定めることができる。

【0096】 推定部11は、これらの予測を行った後に、次式（c-1）で示される拡張カルマンフィルタリングを行い、現時点での状態変数の推定値を算出し、更に、次式（c-2）によって現時点での推定誤差共分散行列の推定値を算出する。

$$X_t = X_t^- + K_t(O_t - f(X_t^-)) \quad \dots(c-1)$$

$$P_t = (I - K_t J_t) P_t^- \quad \dots(c-2)$$

【0097】 式（c-1）において、 $X_t$ は現時点（時刻 $t$ ）の状態変数の推定値であり、 $X_t^-$ は式（b-1）によって得られる現時点（時刻 $t$ ）の状態変数の予測値であり、 $K_t$ は現時点（時刻 $t$ ）でのカルマンゲインであり、 $O_t$ は現時点の処理対象のフレームから実際に抽出された観測ベクトルであり、 $f(X_t^-)$ は式（12）の観測関数に現時点での状態変数の予測値 $X_t^-$ を適用することによって得られる観測ベクトルの推定値である。式（c-2）において、 $P_t$ は現時点で推定誤差共分散行列の推定値、 $P_t^-$ は現時点での推定誤差共分散行列の予測値である。 $J_t$ は観測関数 $f$ のヤコビ行列である。

【0098】 なお、推定部11は、式（c-1）及び式（c-2）による演算を行うために、次式（c-3）の演算によってヤコビ行列 $J_t$ を算出した後、次式（c-4）の演算によってカルマンゲインを算出する。なお、次式（c-4）において、 $C$ は、観測過程において生じるノイズであり、例えばカメラのぶれによって生じるノイズを表している。本実施の形態では、カメラは固定されていて、

このノイズを実質的に無視することができるものとしている。したがって、Cは単位行列となっている。

$$J_t[i, j] = \frac{\partial f_i}{\partial X[j]}(X^{-}_t) \quad \dots(c-3)$$

$$K_t = P^{-}_t J_t^T (J_t P^{-}_t J_t^T + C) \quad \dots(c-4)$$

【0099】 ここで、状態変数Xの各パラメータX[j]に対して、式(c-3)のJ[i, j]は、次式(c-5)となる。

$$J[i, j] = \frac{\partial f_i}{\partial X[j]} = \frac{1}{1 + \beta z_{ci}} \left( \frac{\frac{\partial x_{ci}}{\partial x[j]}}{\frac{\partial y_{ci}}{\partial x[j]}} \right) - \frac{\frac{\partial \beta z_{ci}}{\partial x[j]}}{(1 + \beta z_{ci})^2} \begin{pmatrix} x_{ci} \\ y_{ci} \end{pmatrix} \quad \dots(c-5)$$

【0100】 以下、状態変数Xの各パラメータX[j]について、式(c-5)における $(\frac{\partial x_{ci}}{\partial X[j]}, \frac{\partial y_{ci}}{\partial X[j]}, \frac{\partial z_{ci}}{\partial X[j]})^T$ の演算式を、次式(d-1)～(d-7)に示す。

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial t_x} \\ \frac{\partial y_{ci}}{\partial t_x} \\ \frac{\partial \beta z_{ci}}{\partial t_x} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{\partial x_{ci}}{\partial t_y} \\ \frac{\partial y_{ci}}{\partial t_y} \\ \frac{\partial \beta z_{ci}}{\partial t_y} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{\partial x_{ci}}{\partial t_z} \\ \frac{\partial y_{ci}}{\partial t_z} \\ \frac{\partial \beta z_{ci}}{\partial t_z} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \dots(d-1)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial w} \\ \frac{\partial y_{ci}}{\partial w} \\ \frac{\partial \beta z_{ci}}{\partial w} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot \frac{\partial R(\delta q)}{\partial w} \cdot R_{old} \cdot m_i \quad \dots(d-2)$$

$$\frac{\partial R(\delta q)}{\partial w_x} = \begin{pmatrix} 0 & \frac{2w_y \varepsilon + w_x w_z}{4\varepsilon} & \frac{2w_z \varepsilon - w_x w_y}{4\varepsilon} \\ \frac{2w_y \varepsilon - w_x w_z}{4\varepsilon} & -w_x & \frac{w_x^2}{4\varepsilon} - \varepsilon \\ \frac{2w_z \varepsilon + w_x w_y}{4\varepsilon} & -\frac{w_x^2}{4\varepsilon} + \varepsilon & -w_x \end{pmatrix} \quad \dots(d-2-1)$$

$$\frac{\partial R(\delta q)}{\partial w_y} = \begin{pmatrix} -w_y & \frac{2w_x \varepsilon + w_y w_z}{4\varepsilon} & -\frac{w_y^2}{4\varepsilon} + \varepsilon \\ \frac{2w_x \varepsilon - w_y w_z}{4\varepsilon} & 0 & \frac{2w_z \varepsilon + w_x w_y}{4\varepsilon} \\ \frac{w_y^2}{4\varepsilon} - \varepsilon & \frac{2w_z \varepsilon - w_x w_y}{4\varepsilon} & -w_y \end{pmatrix} \quad \dots(d-2-2)$$

$$\frac{\partial R(\delta q)}{\partial w_z} = \begin{pmatrix} -w_z & \frac{w_z^2}{4\varepsilon} - \varepsilon & \frac{2w_x \varepsilon - w_y w_z}{4\varepsilon} \\ -\frac{w_z^2}{4\varepsilon} + \varepsilon & -w_z & \frac{2w_y \varepsilon + w_x w_z}{4\varepsilon} \\ \frac{2w_x \varepsilon + w_y w_z}{4\varepsilon} & \frac{2w_y \varepsilon - w_x w_z}{4\varepsilon} & 0 \end{pmatrix} \quad \dots(d-2-3)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \beta} \\ \frac{\partial y_{ci}}{\partial \beta} \\ \frac{\partial z_{ci}}{\partial \beta} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} + R_c m_i$$

$$+ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} R_c \cdot \begin{cases} \begin{pmatrix} u_i \alpha_i \\ v_i \alpha_i \\ 0 \end{pmatrix}, i=1, \dots, N \\ \begin{pmatrix} (u_i + \lambda s_i \sin \theta) \alpha_i \\ (v_i + \lambda s_i \sin \phi \cos \theta) \alpha_i \\ \lambda s_i (1 - \cos \phi \cos \theta) \alpha_i \end{pmatrix}, i=N+1, N+2 \end{cases} \quad \dots(d-3)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \alpha_i} \\ \frac{\partial y_{ci}}{\partial \alpha_i} \\ \frac{\partial z_{ci}}{\partial \alpha_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot R_c \cdot \begin{pmatrix} \frac{\partial x_i}{\partial \alpha_i} \\ \frac{\partial y_i}{\partial \alpha_i} \\ \frac{\partial z_i}{\partial \alpha_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot R_c \cdot \begin{pmatrix} u_i \beta \\ v_i \beta \\ 1 \end{pmatrix}, \quad i=1, \dots, N \quad \dots(d-4)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \alpha_i} \\ \frac{\partial y_{ci}}{\partial \alpha_i} \\ \frac{\partial z_{ci}}{\partial \alpha_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot R_c \cdot \begin{pmatrix} \frac{\partial x_i}{\partial \alpha_i} \\ \frac{\partial y_i}{\partial \alpha_i} \\ \frac{\partial z_i}{\partial \alpha_i} \end{pmatrix} \\ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot R_c \cdot \begin{pmatrix} (u_i + \lambda s_i \sin \theta) \beta \\ (v_i + \lambda s_i \sin \phi \cos \theta) \beta \\ 1 + \lambda s_i (1 - \cos \phi \cos \theta) \beta \end{pmatrix}, \quad i = N+1, N+2 \dots (d-5)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \theta} \\ \frac{\partial y_{ci}}{\partial \theta} \\ \frac{\partial z_{ci}}{\partial \theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot R_c \cdot \begin{pmatrix} \lambda s_i \cos \theta (1 + \beta \alpha_i) \\ -\lambda s_i \sin \phi \sin \theta (1 + \beta \alpha_i) \\ \lambda s_i (1 + \cos \phi \sin \theta) (1 + \beta \alpha_i) \end{pmatrix} \dots (d-6)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \phi} \\ \frac{\partial y_{ci}}{\partial \phi} \\ \frac{\partial z_{ci}}{\partial \phi} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \cdot R_c \cdot \begin{pmatrix} 0 \\ \lambda s_i \cos \phi \cos \theta (1 + \beta \alpha_i) \\ \lambda s_i (1 + \sin \phi \cos \theta) (1 + \beta \alpha_i) \end{pmatrix} \dots (d-7)$$

【0101】 推定部11は、フィルタリングによって推定した各パラメータからカメラ座標系 $(X_c, Y_c, Z_c)$ に対する視線座標系 $(X_g, Y_g, Z_g)$ への変換 $T_c^g$  ( $4 \times 4$ の行列)を求める。変換 $T_c^g$ は、式(13)に示すように、カメラ座標系 $(X_c, Y_c, Z_c)$ から顔座標系 $(X_h, Y_h, Z_h)$ への変換 $T_c^h$  ( $4 \times 4$ の行列)と顔座標系 $(X_h, Y_h, Z_h)$ から視線座標系 $(X_g, Y_g, Z_g)$ への変換 $T_h^g$  ( $4 \times 4$ の行列)との積である。変換 $T_c^h$ は、式(11)に示すように、顔の動きに応じて回転行列 $R_c$ 及び並進行列 $t_c$ が設定される。変換 $T_h^g$ は、式(2)に示すように、pベクトル及びqベクトルの各パラメータの推定値を用いて設定される。

$$T_c^g = T_c^h T_h^g = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \dots (13)$$

$$R = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \dots (13-1)$$

【0102】 式(13)において、Rはカメラ座標系( $X_c, Y_c, Z_c$ )に対する瞳の中心点の回転を示す $3 \times 3$ の回転行列であり、tはカメラ座標系( $X_c, Y_c, Z_c$ )に対する瞳の中心点の位置移動を示す $3 \times 1$ の並進行列である(図6参照)。回転行列Rは、次式(e-1)の演算式を用いて算出される。

$$R(\phi, \theta) = R_{x_{e, -\phi}} R_{y_{e, \theta}} \\ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-\phi) & -\sin(-\phi) \\ 0 & \sin(-\phi) & \cos(-\phi) \end{pmatrix} \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \dots (e-1)$$

【0103】 推定部11は、並進行列tで示される位置を始点として回転行列Rの3番目の列ベクトル( $R_{13}, R_{23}, R_{33}$ )<sup>T</sup>で示す方向を視線方向とする。この視線方向は、左右の眼で同じ方向である。ちなみに、回転行列Rの3番目の列ベクトルは、視線座標系( $X_g, Y_g, Z_g$ )におけるZ軸方向である。

【0104】 このように、画像処理装置3は、顔の運動モデルと眼球の運動モデルとを結合したDCCGMを生成し、DCCGMに対する拡張カルマンフィルタによって静的パラメータ及び動的パラメータを推定する。そして、画像処理装置3は、その各パラメータの推定値からカメラ座標系( $X_c, Y_c, Z_c$ )から視線座標系( $X_g, Y_g, Z_g$ )への変換行列を求め、視線方向を求める。

【0105】 以下、画像処理装置3の動作を説明し、併せて本発明の実施の形態に係る画像処理方法を説明する。図7は、本発明の実施の形態に係る画像処理方法の流れを示すフローチャートである。

【0106】 図7に示すように、この画像処理方法では、三次元モデル初期化



部 10 による三次元モデル初期化ステップ S 0 1 と、推定部 11 による運動推定ステップ S 0 2 から構成される。

【0107】 図 8 は、三次元モデル初期化ステップの詳細な処理の流れを示すフローチャートである。三次元モデル初期化ステップ S 0 1 では、まず、三次元モデル初期化部 10 が、観測ベクトル抽出部 12 によって、動画像の初期フレームから、特徴点 ( $i = 1, \dots, N, N+1, N+2$ ) の画像面上における投影座標 ( $u_i, v_i$ ) する (ステップ S 11)。この抽出方法には、上述した固有空間法が用いられる。なお、三次元モデル初期化ステップ S 0 1 にて抽出された投影座標は初期投影座標となる。

【0108】 次いで、三次元モデル初期化部 10 が、 $i = 1, \dots, N$  の顔の特徴点については、式 (1) の  $\alpha_i$ 、 $i = N+1, N+2$  の瞳の中心点については、式 (6) の  $\alpha_i, \phi, \theta$  といったパラメータの初期値を設定することによって、三次元モデルの初期化を行う (ステップ S 0 2)。

【0109】 次いで、推定部 11 による運動推定ステップ S 0 2 が実行される。

図 9 は、運動推定ステップの詳細な処理の流れを示すフローチャートである。図 9 に示すように、運動推定ステップ S 0 2 では、まず、推定部 11 が、式 (b-1) の演算式を実行することによって、現時点  $t$  の状態変数の予測値  $X_t^-$  を算出する (ステップ S 21)。また、推定部 11 が、式 (b-2) の演算式を実行することによって、現時点  $t$  の推定誤差共分散行列の  $P_t^-$  予測値を算出する (ステップ S 22)。

【0110】 次いで、推定部 11 は、現時点  $t$  に対応の処理対象のフレームから、観測ベクトル  $O_i$  ( $i = 1, \dots, N, N+1, N+2$ ) を抽出する (ステップ S 23)。次いで、推定部 11 は、拡張カルマンフィルタリングによって現時点  $t$  の状態変数の推定値  $X_t$  を算出する (ステップ S 24)。図 10 は、拡張カルマンフィルタを用いる処理の詳細な流れを示すフローチャートである。

【0111】 図 10 に示すように、ステップ S 24 において、推定部 11 は、

まず、式 (c-3) の演算式を実行することによって、観測行列  $f$  のヤコビ行列  $J_t$  を算出する (ステップ S 3 1)。次いで、推定部 1 1 は、式 (c-4) の演算式を実行することによって、カルマンゲイン  $K_t$  を算出する (ステップ S 3 2)。

【0 1 1 2】 そして、推定部 1 1 は、現時点のモデル座標を更新する (ステップ S 3 3)。ステップ S 3 3 では、現時点  $t$  の状態変数の予測値  $X_t$  に含まれているパラメータ及び初期投影座標が式 (7-1-1) の演算式に適用されることによって、 $i = 1, \dots, N$  の特徴点のモデル座標が算出される。また、現時点  $t$  の状態変数の予測値  $X_t$  に含まれているパラメータ及び初期投影座標が式 (6) の演算式に適用されることによって、 $i = N + 1, N + 2$  の特徴点 (瞳の中心点) のモデル座標が算出される。

【0 1 1 3】 次いで、推定部 1 1 は、ステップ S 3 3 にて抽出されたモデル座標を式 (1 1) の運動演算式に適用することによって、現時点  $t$  での  $i = 1, \dots, N, N + 1, N + 2$  の特徴点の座標の推定値を算出する (ステップ S 3 4)。

【0 1 1 4】 次いで、推定部 1 1 は、ステップ S 3 4 にて算出した特徴点の座標の推定値を式 (1 2) の観測関数に適用することによって、現時点  $t$  での特徴点の投影座標の推定値を算出する (ステップ S 3 5)。ステップ S 3 5 で算出された投影座標は、現時点  $t$  での観測ベクトル  $O_t$  の推定値である。

【0 1 1 5】 次いで、推定部 1 1 は、式 (c-1) のフィルタ方程式の演算を実行することによって、現時点  $t$  の状態変数の推定値  $X_t$  を算出する (ステップ S 3 6)。そして、推定部 1 1 は、式 (c-2) の演算式を実行することによって、現時点  $t$  の推定誤差共分散行列の推定値  $P_t$  を算出する (ステップ S 3 7)。

【0 1 1 6】 図 9 に戻り、次に、推定部 1 1 は、状態変数の推定値  $X_t$  に基づき、運動を表す力値を出力する (ステップ S 2 5)。第 1 の実施の形態では、ステップ S 2 5 において、現時点  $t$  の状態変数の推定値  $X_t$  の  $\phi, \theta$  が式 (e-1) の演算式に適用されることによって、回転行列  $R$  が算出される。そして、回転行列  $R$  の第 3 番目の列ベクトル、及び並進行列  $t$  が視線方向の運動を表す出力値として

出力される。

【0117】 次に、推定部11は、全フレームの処理が終了したか否かをテストし（ステップS26）、全フレームの処理が終了していない場合（NOの場合）には、現時点tの時刻を更新（ $t = t + 1$ ）し（ステップS27）、ステップS21からの処理を繰り返す。一方、推定部11は全フレームの処理が終了している場合（YESの場合）には、処理を終了する。

【0118】 以下、コンピュータを本発明の画像処理装置として動作させる画像処理プログラム及び、当該画像処理プログラムが記録された記録媒体の実施の形態について説明する。図11は、本発明の実施の形態に係る画像処理プログラムの構成を、記録媒体と共に示す図である。

【0119】 図11に示すように、画像処理プログラム50は、記録媒体100に格納されて提供される。記録媒体100としては、フロッピーディスク、CD-ROM、DVD、あるいはROM等の記録媒体あるいは半導体メモリ等が例示される。

【0120】 図12は、記録媒体に記憶されたプログラムを実行するためのコンピュータのハードウェア構成を示す図であり、図13は、記録媒体に記憶されたプログラムを実行するためのコンピュータの斜視図である。図12に示すように、コンピュータ110は、フロッピーディスクドライブ装置、CD-ROMドライブ装置、DVDドライブ装置等の読取装置112と、オペレーティングシステムを常駐させた作業用メモリ（RAM）114と、記録媒体100に記憶されたプログラムを記憶するメモリ116と、ディスプレイといった表示装置118と、入力装置であるマウス120及びキーボード122と、データ等の送受を行うための通信装置124と、プログラムの実行を制御するCPU126とを備えている。コンピュータ110は、記録媒体100が読取装置112に挿入されると、読取装置112から記録媒体100に格納された画像処理プログラム50にアクセス可能になり、当該画像処理プログラム50によって、画像処理装置3と

して動作することが可能になる。

【0121】 図13に示すように、画像処理プログラム50は、搬送波に重畳されたコンピュータデータ信号130としてネットワークを介して提供されるものであってもよい。この場合、コンピュータ110は、通信装置124によって受信した画像処理プログラム50をメモリ116に格納し、当該画像処理プログラム50を実行することができる。

【0122】 図11に示すように、画像処理プログラム50は、処理を統括するメインモジュール52と、観測ベクトル抽出モジュール54と、三次元モデル初期化モジュール56と、運動推定モジュール58とによって構成されている。

【0123】 観測ベクトル抽出モジュール54は、上述した観測ベクトル抽出部12と同様の機能をコンピュータに実行させる。三次元モデル初期化モジュール56は上述した三次元モデル初期化ステップS01を、運動推定モジュール58は上述した運動推定ステップS02をコンピュータに実行させる。三次元モデル初期化モジュール56は、及び運動推定モジュール58は、その処理の中で、適宜観測ベクトル抽出モジュール54を呼び出す。

【0124】 以下、画像処理システム1の有効性の確認のために行われた実験の結果を説明する。図14A及び図14Bに、画像処理システム1によって実際に推定した視線方向を顔画像に描いた一例を示す。この例では、 $320 \times 240$ の画像サイズを有する動画像が用いられている。図14A及び図14Bでは、黒色の矢印で顔姿勢の方向が示されており、白色の矢印で視線方向が示されており、大きい白色の円で瞳の外周が示されており、多数の小さい白色の円で顔の特徴点が示されている。図14Aの例では、顔姿勢が左方向であり、視線が顔姿勢より中心よりの左方向である。図14Bの例では、顔姿勢が右方向であり、視線が左方向である。

【0125】 画像処理システム1（画像処理装置3）によれば、実際に起こりうる瞳の動きを数学式でモデリングするとともに静的パラメータや動的パラメー

タを直接フィルタリングしているので、視線方向を高精度に推定することができる。また、画像処理システム 1 では、単眼カメラで制限のない顔の動きに対して視線の推定が可能である。さらに、画像処理システム 1 では、最も一般的な透視投影変換をカメラモデルに使用しているので、応用範囲が広い。また、画像処理システム 1 では、非接触式で視線の推定が可能であり、高価なカメラシステムや装置を必要としない。

【0126】 [第2の実施の形態]

【0127】 第2の実施の形態に係る画像処理システムを説明する前に、図15A、図15B、及び図15Cを参照して、顔の唇の運動について説明しておく。図15A、図15B、及び図15Cは、顔の唇の各特徴点を示す顔画像を模式的に示しており、図15Aが顔の動き及び唇の動きが初期状態の場合を示し、図15Bが唇のみが動いた場合を示し、図15Cが図15Bの唇の動きに顔の動きを伴った場合を示している

【0128】 図15Aの顔画像では、顔が真正面を向いており、唇が閉じた状態である初期状態を示す。唇には、左右端及び中央の上下端に特徴点（十印参照）が設定されている。図15Bの顔画像では、顔が真正面を向いたままで、唇を開いた状態を示す。このように、唇の動きによって、唇の特徴点に変化する。図15Cの顔画像では、顔が斜め前方を向き、唇を開いた状態を示す。このように、顔の動きを伴うことによって、唇の特徴点が更に変化する。つまり、顔の運動と唇の三次元形状の変化によって、これら3つの画像上の特徴点の動きは互いに異なる。

【0129】 このように、唇に設定された特徴点の動きから唇の運動情報を求める場合、顔画像上の特徴点には顔の動きと唇の動きとが混在する。そのため、顔の運動と唇の三次元形状の構造が判らないと、non-rigidな変化をする唇の運動情報を求めることができない。そこで、第2の実施の形態では、顔の三次元構造を、rigid構造を表すためのrigidパラメータとnon-r

rigidな変化を表すためのnon-rigidパラメータによって定義する。rigid構造は、顔に表情の変化がない場合（顔の三次元形状にnon-rigidな変化がない場合）の顔の三次元構造である。non-rigidな変化は、顔の表情の変化による顔の三次元構造の変化である。なお、顔には、non-rigidな変化をする箇所としては、唇だけでなく、目、鼻、頬等の様々な箇所がある。

【0130】 以下、図1を参照して、第2の実施の形態に係る画像処理システム21について説明する。なお、第2の実施の形態では、第1の実施の形態に係る画像処理システムと同様の構成については同一の符号を付し、その説明を省略する。

【0131】 画像処理システム21は、透視投影変換のカメラモデルを用い、非接触で顔の各部（唇等）の運動の推定を行う。画像処理システム21では、画像取得装置2が画像処理装置23に接続され、画像処理装置23において画像取得装置2で取得した動画像データから顔の各部の運動の最適推定を行う。画像処理システム21でも、必要に応じて、モニタ等の出力手段が画像処理装置23に接続される。

【0132】 図1、図16及び図17を参照して、画像処理装置23について説明する。図16は、モデル座標系を示す図である。図17は、顔が並進及び回転運動した場合のモデル座標系を示す図である。

【0133】 画像処理装置23は、第1の実施の形態に係る画像処理装置3と同様の装置であり、観測ベクトル抽出部12、三次元モデル初期化部30及び推定部31を備えるが、各部30、31で行う処理が異なる。なお、画像処理装置23では、 $N$ 個（ $i = 1, 2, \dots, N$ ）の特徴点を取り扱う。また、画像処理装置23では、画像処理を行う上での基準の座標系であるモデル座標系（ $X_0, Y_0, Z_0$ ）が定義されている。

【0134】 三次元モデル初期化部30について説明する。三次元モデル初期

化部 30 は、初期フレームに基づいて、顔の特徴点の三次元構造を定義する。初期フレームでは、顔が固定であり、顔座標系 ( $X_h, Y_h, Z_h$ ) がカメラ座標系 ( $X_c, Y_c, Z_c$ ) に一致している。なお、第 2 の実施の形態では、このカメラ座標系 ( $X_c, Y_c, Z_c$ ) をモデル座標系として定義している。

- 5 【0135】 三次元モデル初期化部 30 は、初期フレームの画像から  $N$  個の顔の特徴点 ( $x_i, y_i, z_i$ ) ( $i = 1, 2, \dots, N$ ) の投影座標 ( $u_i, v_i$ ) を抽出する。三次元モデル初期化部 30 は、抽出した投影座標 ( $u_i, v_i$ ) を用いて、式 (14) に示すモデル座標演算式によって顔の特徴点のモデル座標 ( $x_i, y_i, z_i$ ) を定義し、顔の三次元構造を定義している。すなわち、三次元モデル初期化部 30 は、モデル座標を定義する式 (14) におけるパラメータの初期値を設定する。顔の特徴点 ( $x_i, y_i, z_i$ ) は、モデル座標上の位置である。

$$\vec{m}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} (u_i + \phi_i) & (1 + \alpha_i \beta) \\ (v_i + \zeta_i) & (1 + \alpha_i \beta) \\ \alpha_i + \delta_i \end{pmatrix} \cdot \dots \quad (14)$$

- 15 【0136】 図 16 は、初期フレームで定義するモデル座標系を示している。図 16 において、 $(0, 0, -f)$  が CCD カメラの焦点であり、焦点距離  $f$  離れたところに画像面が位置している。モデル座標系の原点  $(0, 0, 0)$  は、画像面上にある。画像面に顔の特徴点 ( $x_i, y_i, z_i$ ) が投影されると、観測点 ( $u_i, v_i, 0$ ) となる。モデル座標系の  $Z$  軸方向において、原点  $(0, 0, 0)$  から特徴点 ( $x_i, y_i, z_i$ ) までの奥行きが  $\alpha_i$  となる。モデル座標系では、焦点距離  $f$  と奥行き  $\alpha$  とに分離されている。

- 20 【0137】 式 (14) は、顔の特徴点のモデル座標を演算するためのモデル座標演算式を示している。 $(u_i, v_i, 0)$  は、顔の特徴点を CCD 上に投影した初期投影座標 (画像座標) であり、初期フレームで求めた後はその値に固定される。 $\alpha_i$  は、顔の特徴点 ( $x_i, y_i, z_i$ ) から CCD までの距離であり、特徴点の奥行きを示している。 $\alpha_i$  は、決まった値を有しているが、未知である。 $\psi_i$

は、初期フレームから現フレームまでの $u_i$ の変化量であり、画像上の特徴点のピクセル移動量を示す。 $\xi_i$ は、初期フレームから現フレームまでの $v_i$ の変化量であり、画像上の特徴点のピクセル移動量を示す。 $\delta_i$ は、初期フレームから現フレームまでの $\alpha_i$ の変化量であり、奥行き $\alpha_i$ の移動量を示す。 $\psi_i$ 、 $\xi_i$ 、 $\delta_i$ は、顔の表情の変化に応じて変化する値であり、未知である。 $\beta$ は、CCDカメラの焦点距離 $f$ の逆数( $f^{-1}$ )であり、CCDカメラによって変わり、未知である。

【0138】 各特徴点の奥行き $\alpha_i$ は、顔の各特徴点の三次元構造のrigid構造を表し、顔の三次元形状に対応する。また、 $\psi_i$ 、 $\xi_i$ 、 $\delta_i$ は、顔の各特徴点の三次元構造のnon-rigidな変化を表す、唇の動き等の顔の表情の変化に対応する。奥行き $\alpha_i$ の移動量 $\delta_i$ は、顔の表情が変化した場合(すなわち、顔の三次元形状がnon-rigidな変化をした場合)、その値は殆ど変わらないので、無視することができる。したがって、 $\delta_i$ を0とし、初期フレームから現フレームまでの画像を用いて、モデル座標系上における初期フレームでの特徴点からのピクセル移動量( $\psi_i$ 、 $\xi_i$ )を帰納的に推定することにより、顔の各特徴点でのnon-rigidな変化が求められることになる。

【0139】 1つの三次元特徴点に対しては、一般にrigid構造とそのnon-rigidな変化を表すためには6個のパラメータが必要であるが、ここでは3つのパラメータでそれを可能としている。すなわち、rigid構造を表すためのrigidパラメータ( $\alpha_i$ )とnon-rigidな変化を表すためのnon-rigidパラメータ( $\psi_i$ 、 $\xi_i$ )の3つのパラメータで三次元特徴点を表すことができる。

【0140】 推定部31について説明する。推定部31では、顔の運動モデルを定義し、顔の運動モデルに対して拡張カルマンフィルタリングを施し、顔の各部の運動を推定する。

【0141】 推定部31は、顔の運動モデルとして、回転運動と並進運動とに



よって表現される運動モデルを想定している。顔の動きをモデル座標系 ( $X_c$ ,  $Y_c$ ,  $Z_c$ ) に対して回転と並進移動で表すと、モデル座標から顔の特徴点 ( $x_{ci}$ ,  $y_{ci}$ ,  $\beta z_{ci}$ ) は、次式 (15-1) によって推定される (図17参照)。

$$\begin{pmatrix} x_{ci} \\ y_{ci} \\ \beta z_{ci} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} + R_c \bar{m}_i \quad \dots (15-1)$$

5 【0142】 式 (15-1) において、( $t_x$ ,  $t_y$ ,  $t_z$ ) はモデル座標系 ( $X_i$ ,  $Y_i$ ,  $Z_i$ ) からの  $3 \times 1$  の並進行列  $t_c$  であり、 $R_c$  はモデル座標系 ( $X_c$ ,  $Y_c$ ,  $Z_c$ ) からの  $3 \times 3$  の回転行列であり、 $m_i$  ベクトル、すなわちモデル座標は式 (14) で表すことができる。顔の動きには、制限がなく、並進行列  $t_c$  と回転行列  $R_c$  によってあらゆる動きを表すことができる。

10 【0143】 この顔の特徴点の座標の推定値 ( $x_{ci}$ ,  $y_{ci}$ ,  $\beta z_{ci}$ ) を画像面 (CCD) 上に投影した  $o_i$  ベクトル (観測ベクトル) の推定値は、式 (15-1) の演算式によって算出される顔の特徴点の座標の推定値 ( $x_{ci}$ ,  $y_{ci}$ ,  $\beta z_{ci}$ ) から次式 (15-2) によって算出される (図17参照)。

$$\bar{o}_i = \begin{pmatrix} \frac{x_{ci}}{1 + \beta z_{ci}} \\ \frac{y_{ci}}{1 + \beta z_{ci}} \end{pmatrix} \quad \dots (15-2)$$

15 ここで、全ての顔の特徴点 ( $i = 1, 2, \dots, N$ ) に対する観測ベクトルである  $O$  ベクトルは、次式 (15-3) で表すことができる。このように、推定部31は、顔の特徴点に対する観測モデルを表す。

$$\bar{O} = \begin{pmatrix} \bar{o}_1 \\ \bar{o}_2 \\ \vdots \\ \bar{o}_N \end{pmatrix} \dots (15-3)$$

【0144】 推定部31では、この顔の運動モデルの状態空間表現（状態遷移モデル）の状態変数であるXベクトルが次式（16-1）で定義されている。Xベクトルは、顔の姿勢とrigidパラメータ及びnon-rigidパラメータで構成される。さらに、推定部31では、顔の運動モデルの状態空間表現が式（16-2）で定義されている。

$$\bar{X} = \begin{pmatrix} t_x \\ t_y \\ \beta t_z \\ w_x \\ w_y \\ w_z \\ \beta \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \\ \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \\ \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_N \end{pmatrix} \dots (16-1)$$

$$\bar{X}_t = A^{(1)} \bar{X}_{t-1} + \mu_t^{(1)} \dots (16-2)$$

【0145】 式（16-1）において、 $(t_x, t_y, \beta t_z)$  は顔のモデル座標

系  $(X_c, Y_c, Z_c)$  における並進移動量であり、 $(w_x, w_y, w_z)$  は前フレームからのモデル座標系  $(X_c, Y_c, Z_c)$  における各軸の回転量であり、 $\beta$  は CD カメラの焦点距離の逆数であり、 $\alpha_1 \sim \alpha_N$  は顔の特徴点の各奥行きであり、 $\psi_1 \sim \psi_N$  は画像上の特徴点座標である  $u_1 \sim u_N$  の初期フレームから現フレームまでの各変化量であり、 $\xi_1 \sim \xi_N$  は画像上の特徴点座標である  $v_1 \sim v_N$  の初期フレームから現フレームまでの各変化量である。式 (16-2) において、下付き文字の  $t$  は現時点（現フレームに対応する時刻）を示し、下付き文字の  $t-1$  は一時刻前（前フレームに対応する時刻）を示し、 $A^{(1)}$  は現状態と一時刻前の状態との顔の運動特性を表す状態遷移行列であり、 $\mu_t^{(1)}$  は現状態と一時刻前の状態とのノイズを表す。特に、 $\mu_t^{(1)}$  は、ゼロの平均と分散共分散行列で構成されるガウシアン分布のランダムベクトルである。

【0146】 推定部 31 では、顔の運動モデルの状態変数である  $X_t$  ベクトルと観測ベクトルである  $O_t$  ベクトルとの関係が次式 (17) で定義されている。 $O_t$  ベクトルは、 $N$  個分の  $o_i$  ベクトルを含んでいる。式 (17) において、 $f()$  は観測関数であり、 $v_t$  はノイズを表し、ゼロの平均とノイズの共分散行列  $R^{(1)}$  とで構成されるランダムベクトルである。ここで、 $X_t$  ベクトルを変数とする観測関数  $f()$  は、各特徴点の座標  $(x_{ci}, y_{ci}, z_{ci})$  の推定値を用いる観測関数  $f_i()$  によって、次式 (17-1) のように表現される。

$$\bar{O}_t = \begin{pmatrix} \bar{o}_1 \\ \bar{o}_2 \\ \vdots \\ \bar{o}_N \end{pmatrix} = f(\bar{X}_t) + v_t \quad \dots (17)$$

$$f(\bar{X}_t) = \begin{pmatrix} f_1(\bar{X}_t) \\ \vdots \\ f_N(\bar{X}_t) \end{pmatrix} \cdots (17-1)$$

【0147】 さらに、推定部31では、式(17)で示す観測ベクトル $O_t$ から状態変数 $X_t$ を推定するために、顔の各特徴点の座標の推定値( $x_{ci}$ ,  $y_{ci}$ ,  $z_{ci}$ )を画像面(CCD)上に投影した $o_i$ ベクトル(観測ベクトル)の推定値が次式(17-2)で定義されている

$$\bar{o}_i = \begin{pmatrix} \frac{x_{ci}}{1 + \beta z_{ci}} \\ \frac{y_{ci}}{1 + \beta z_{ci}} \end{pmatrix} = f_i(\bar{X}_t) \cdots (17-2)$$

【0148】 式(17)から、状態変数である $X_t$ ベクトルと観測ベクトルである $O_t$ ベクトルとの関係は非線形であることが判る。そこで、推定部31は、式(17)の観測ベクトル $O_t$ に対して拡張カルマンフィルタリングを施し、式(16-1)の顔の運動モデルの状態変数である $X_t$ ベクトルの各パラメータ( $t_x$ ,  $t_y$ ,  $\beta t_z$ ,  $w_x$ ,  $w_y$ ,  $w_z$ ,  $\alpha_1 \sim \alpha_N$ ,  $\beta$ ,  $\psi_1 \sim \psi_N$ ,  $\xi_1 \sim \xi_N$ )を推定する。この観測過程では、式(15-2)の観測ベクトル $o_i$ 及び式(15-1)の三次元構造を用いる。

【0149】 推定部31は、推定したパラメータを用いて、顔の並進及び回転運動を考慮しながら、顔の三次元構造と初期フレームで設定した特徴点のnon-rigidな変化を推定し、顔の各部の運動情報を求める。

【0150】 このように、画像処理装置23では、顔の運動モデルに対する拡張カルマンフィルタによって状態変数である $X_t$ ベクトルの各パラメータを推定し、顔の並進及び回転の運動、顔の各特徴点でのrigid構造及び顔の各特徴点でのnon-rigid変化を推定する。そして、画像処理装置23では、各

パラメータの推定値から、顔の動きに対応しながら *non-rigid* な変形を伴う顔の各部の運動を求めている。

【0151】 以下、この画像処理装置23の動作及び画像処理装置23で実行される画像処理方法を説明する。ここでは、上述の図7～図10を用いて、第1の実施の形態と異なる動作についてのみ説明する。

【0152】 第2の実施の形態の画像処理方法も、三次元モデル初期化ステップS01及び運動推定ステップS02によって構成されている。三次元モデル初期化ステップS01では、三次元モデル初期化部30が、ステップS11において、N個の顔の特徴点の初期投影座標を抽出する。また、ステップS12において、三次元モデル初期化部30が、上記の $\alpha_1 \sim \alpha_N$ 、 $\psi_1 \sim \psi_N$ 、 $\xi_1 \sim \xi_N$ といったパラメータの初期値を設定することによって、三次元モデルを初期化する。

【0153】 運動推定ステップS02では、推定部31が、第1の実施の形態と同様のステップS21からステップS23の処理を行い、ステップS24において、推定部31は、観測関数fのヤコビ行列 $J_t$ を算出する(ステップS31)。ヤコビ行列は、式(c-3)及び式(c-5)によって求められるが、第2の実施の形態は第1の実施の形態と状態変数Xが異なるので、上述の式(d-1)～(d-4)、次式(d-8)及び(d-9)の演算式が実行される。

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \zeta_i} \\ \frac{\partial y_{ci}}{\partial \zeta_i} \\ \frac{\partial \beta z_{ci}}{\partial \zeta_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} R_c \begin{pmatrix} \frac{\partial x_i}{\partial \zeta_i} \\ \frac{\partial y_i}{\partial \zeta_i} \\ \frac{\partial z_i}{\partial \zeta_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} R_c \begin{pmatrix} 1 + \alpha_i \beta \\ 0 \\ 0 \end{pmatrix} \quad \dots (d-8)$$

$$\begin{pmatrix} \frac{\partial x_{ci}}{\partial \psi_i} \\ \frac{\partial y_{ci}}{\partial \psi_i} \\ \frac{\partial z_{ci}}{\partial \psi_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} R_c \begin{pmatrix} \frac{\partial x_i}{\partial \psi_i} \\ \frac{\partial y_i}{\partial \psi_i} \\ \frac{\partial z_i}{\partial \psi_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix} R_c \begin{pmatrix} 0 \\ 1 + \alpha_i \beta \\ 0 \end{pmatrix} \quad \dots (d-9)$$

【0154】 次いで、推定部31は、現時点 $t$ の状態変数の予測値 $X_t^-$ に含まれているパラメータを式(14)の演算式に適用することによって、現時点 $t$ のモデル座標 $m_i$ を更新し(ステップS33)、式(15-1)の演算式を実行することによって、現時点 $t$ の特徴点の座標の推定値を算出する(ステップS34)。

【0155】 次いで、推定部31は、式(17-2)の観測関数 $f$ に現時点 $t$ の特徴点の座標の推定値を代入することによって、現時点 $t$ での観測ベクトル $O_i$ の推定値を算出する(ステップS35)。

【0156】 そして、推定部31は、第1の実施の形態と同様にステップS36及びステップS37の処理を実行し、現時点 $t$ の状態変数の推定値 $X_t$ を算出する。次いで、推定部31は、ステップS25において、現時点 $t$ の状態変数の推定値 $X_t$ のパラメータをそのまま、顔の各部の運動を表す出力値として出力する。次いで、推定部31は、第1の実施形態と同様のステップS26を実行し、ステップS26のテストがNOの場合には、ステップS27へ移行し、YESの場合には処理を終了する。

【0157】 以下、第2の実施の形態の画像処理装置23としてコンピュータを動作させる画像処理プログラム及び当該画像処理プログラムを記録した記録媒体について説明する。図11に示すように、第2の実施の形態の画像処理プログラム50も第1の実施の形態の画像処理プログラムと同様の構成を有し、記録媒体100についても同様である。第2の実施の形態の画像処理プログラム50では、三次元モデル初期化モジュール56が、第2の実施の形態の三次元モデル初期化ステップS01をコンピュータに実行させ、運動推定モジュール58が、第

2の実施の形態の運動推定ステップS02をコンピュータに実行させる。

【0158】 以上説明したように、画像処理システム21（画像処理装置23）によれば、顔の三次元構造をrigidパラメータ及びnon-rigidパラメータでモデリングするとともにrigidパラメータやnon-rigidパラメータを直接フィルタリングしているので、顔の表情が変化した場合でも、顔の各部の運動を高精度に推定することができる。特に、画像処理システム21によれば、顔の動きに関わらず、顔の各部のnon-rigidな運動情報を求めることができる。また、画像処理システム21は、単眼カメラで制限のない顔の動きに対して顔の各部の運動の推定が可能である。さらに、画像処理システム21は、最も一般的な透視投影変換をカメラモデルに使用しているので、応用範囲が広い。また、画像処理システム21では、非接触式で顔の各部の運動の推定が可能であり、高価なカメラシステムや装置を必要としない。

【0159】 以上、本発明に係る実施の形態について説明したが、本発明は上記実施の形態に限定されることなく様々な形態で実施される。

【0160】 例えば、本実施の形態ではコンピュータでアプリケーションプログラムを実行することによって画像処理装置を構成したが、視線や顔の各部の運動を推定するための専用の画像処理装置で構成してもよいし、あるいは、人の視線情報や顔の各部の運動情報を必要とする装置（例えば、ロボット、自動車の安全装置）に組み込まれる構成でもよい。

【0161】 また、本実施の形態では視線を推定するための画像処理装置とnon-rigidな変化を伴う顔の各部の運動を推定するための画像処理装置を別々に構成したが、視線とnon-rigidな変化を伴う顔の各部の運動を推定するための画像処理装置を一体で構成してもよい。

【0162】 また、本実施の形態では状態変数を推定するために拡張カルマンフィルタを用いたが、他のフィルタやフィルタ以外の手段等によって推定してもよい。

【0163】 また、第2の実施の形態では三次元物体として顔に適用したが、三次元物体全体が並進及び回転運動をするとともに三次元物体の各部がnon-rigidな変化する三次元物体に適用可能であり、例えば、心臓に適用可能である。

## 5 産業上の利用可能性

【0164】 本発明によれば、三次元物体の運動を高精度に求めることができる。特に、視線を高精度に求めることができる。また、non-rigidな変化を伴う三次元物体の運動を高精度に求めることができる。



## 請求の範囲

1. 単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を推定する画像処理装置であって、

5 前記動画像のフレーム各々から、前記所定の特徴点の画像面への投影座標を抽出する観測ベクトル抽出手段と、

第1のパラメータと第2のパラメータと初期投影座標とに基づいて前記所定の特徴点のモデル座標を算出するためのモデル座標演算式における該初期投影座標を前記観測ベクトル抽出手段によって前記動画像の初期フレームから抽出する三次元モデル初期化手段と、

10 前記初期フレームとは異なる前記動画像の処理対象フレームが撮影された撮影時点での前記所定の特徴点の座標を前記モデル座標から算出するための運動演算式における第3のパラメータと、前記第1のパラメータと、前記第2のパラメータとを含む状態変数の推定値を算出し、該状態変数の推定値に含まれる第2のパラメータに基づいて前記所定の特徴点の運動に関する出力値を出力する運動推定手段と、

15 を備え、

前記モデル座標演算式は前記単眼カメラの逆投影に基づいており、前記第1のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存しないパラメータであり、前記第2のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存するパラメータであり、

20 前記運動推定手段は、

前記処理対象フレームが撮影された撮影時点における前記状態変数の予測値を状態遷移モデルに基づいて算出し、

25 前記初期投影座標と、前記状態変数の予測値に含まれている前記第1のパラメータと前記第2のパラメータとを、前記モデル座標演算式に適用することによって、前記撮影時点におけるモデル座標の推定値を算出し、

前記状態変数の予測値に含まれている前記第3のパラメータと前記モデル座標の推定値とを前記運動演算式に適用することによって、前記撮影時点における前記所定の特徴点の座標の推定値を算出し、

5 前記所定の特徴点の座標の推定値を、前記単眼カメラの観測モデルに基づく観測関数に適用することによって、前記所定の特徴点の観測ベクトルの推定値を算出し、

前記観測ベクトル抽出手段によって前記処理対象フレームから前記所定の特徴点の投影座標を観測ベクトルとして抽出し、

10 抽出した前記観測ベクトルと前記観測ベクトルの推定値とを用いて前記状態変数の予測値をフィルタリングすることによって、前記撮像時点における前記状態変数の推定値を算出する、  
画像処理装置。

2. 前記第1のパラメータは、特定の値に収束する静的パラメータであり、  
前記第2のパラメータは、前記所定の特徴点を含む部分の運動によって変化する動的パラメータである、請求項1記載の画像処理装置。  
15

3. 前記静的パラメータは、前記画像面から前記所定の特徴点への奥行きである、請求項2記載の画像処理装置。

4. 前記動的パラメータは、前記所定の特徴点を含む部分の回転運動を特定するための回転パラメータである、請求項2又は3記載の画像処理装置。

20 5. 前記回転パラメータは、前記所定の特徴点を含む部分の中心を原点とする座標系における2つの座標軸に対して、該原点から該所定の特徴点へのベクトルが各々なす角度である、請求項4に記載の画像処理装置。

6. 前記第1のパラメータは、`rigid`パラメータであり、  
前記第2のパラメータは、`non-rigid`パラメータである、請求項1に  
25 記載の画像処理装置。

7. 前記`rigid`パラメータは、前記画像面から前記モデル座標までの奥

行きである、請求項 6 に記載の画像処理装置。

8. 前記 `non-rigid` パラメータは、前記所定の特徴点を含む部分の運動による前記所定の特徴点の位置変化に関する変化量である、請求項 6 又は 7 に記載の画像処理装置。

5 9. 前記運動モデルは、前記三次元物体の回転及び並進運動に基づいており、前記第 3 のパラメータは、前記三次元物体の並進量を特定するための並進パラメータ及び該三次元物体の回転量を特定するための回転パラメータである、請求項 1 ～ 8 の何れか 1 項に記載の画像処理装置。

10 10. 前記運動推定手段は、前記フィルタリングとして拡張カルマンフィルタリングを適用する、請求項 1 ～ 9 の何れか 1 項に記載の画像処理装置。

11. 単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を推定する画像処理方法であって、

15 第 1 のパラメータと第 2 のパラメータと初期投影座標とに基づいて前記所定の特徴点のモデル座標を算出するためのモデル座標演算式における該初期投影座標を前記動画像の初期フレームから抽出する三次元モデル初期化ステップと、

前記初期フレームとは異なる前記動画像の処理対象フレームが撮影された撮影時点での前記所定の特徴点の座標を前記モデル座標から算出するための運動演算式における第 3 のパラメータと、前記第 1 のパラメータと、前記第 2 のパラメータとを含む状態変数の推定値を算出し、該状態変数の推定値に含まれる第 2 のパラメータに基づいて前記所定の特徴点の運動に関する出力値を出力する運動推定ステップと、

20 を含み、

前記モデル座標演算式は前記単眼カメラの逆投影に基づいており、前記第 1 のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存しないパラメータであり、前記第 2 のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存するパラメータであり、

25

前記運動推定ステップにおいて、

前記処理対象フレームが撮影された撮影時点における前記状態変数の予測値を状態遷移モデルに基づいて算出し、

5 前記初期投影座標と、前記状態変数の予測値に含まれている前記第 1 のパラメータと前記第 2 のパラメータとを、前記モデル座標演算式に適用することによって、前記撮影時点におけるモデル座標の推定値を算出し、

前記状態変数の予測値に含まれている前記第 3 のパラメータと前記モデル座標の推定値とを前記運動演算式に適用することによって、前記撮影時点における前記所定の特徴点の座標の推定値を算出し、

10 前記所定の特徴点の座標の推定値を、前記単眼カメラの観測モデルに基づく観測関数に適用することによって、前記所定の特徴点の観測ベクトルの推定値を算出し、

前記処理対象フレームから前記所定の特徴点の投影座標を観測ベクトルとして抽出し、

15 抽出した前記観測ベクトルと前記観測ベクトルの推定値とを用いて前記状態変数の予測値をフィルタリングすることによって、前記撮像時点における前記状態変数の推定値を算出する、

画像処理方法。

1 2. 前記第 1 のパラメータは、特定の値に収束する静的パラメータであり、

20 前記第 2 のパラメータは、前記所定の特徴点を含む部分の運動によって変化する動的パラメータである、請求項 1 1 記載の画像処理方法。

1 3. 前記静的パラメータは、前記画像面から前記所定の特徴点への奥行きである、請求項 1 2 記載の画像処理方法。

25 1 4. 前記動的パラメータは、前記所定の特徴点を含む部分の回転運動を特定するための回転パラメータである、請求項 1 2 又は 1 3 記載の画像処理方法。

1 5. 前記回転パラメータは、前記所定の特徴点を含む部分の中心を原点と

する座標系における2つの座標軸に対して、該原点から該所定の特徴点へのベクトルが各々なす角度である、請求項14に記載の画像処理方法。

16. 前記第1のパラメータは、*rigid*パラメータであり、

5 前記第2のパラメータは、*non-rigid*パラメータである、請求項11に記載の画像処理方法。

17. 前記*rigid*パラメータは、前記画像面から前記モデル座標までの奥行きである、請求項16に記載の画像処理方法。

18. 前記*non-rigid*パラメータは、前記所定の特徴点を含む部分  
10 の運動による前記所定の特徴点の位置変化に関する変化量である、請求項16又は17に記載の画像処理方法。

19. 前記運動モデルは、前記三次元物体の回転及び並進運動に基づいており、

前記第3のパラメータは、前記三次元物体の並進量を特定するための並進パラメータ及び該三次元物体の回転量を特定するための回転パラメータである、  
15 請求項11～18の何れか1項に記載の画像処理方法。

20. 前記フィルタリングとして拡張カルマンフィルタリングを適用する、請求項11～19の何れか1項に記載の画像処理方法。

21. 単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を推定するようにコンピュータを動作させる画像処理プログラムであって、該コンピュータに、  
20

第1のパラメータと第2のパラメータと初期投影座標とに基づいて前記所定の特徴点のモデル座標を算出するためのモデル座標演算式における該初期投影座標を前記動画像の初期フレームから抽出する三次元モデル初期化ステップと、

前記初期フレームとは異なる前記動画像の処理対象フレームが撮影された撮影  
25 時点での前記所定の特徴点の座標を前記モデル座標から算出するための運動演算式における第3のパラメータと、前記第1のパラメータと、前記第2のパラメータ

タとを含む状態変数の推定値を算出し、該状態変数の推定値に含まれる第2のパラメータに基づいて前記所定の特徴点の運動に関する出力値を出力する運動推定ステップと、

を実行させ、

- 5       前記モデル座標演算式は前記単眼カメラの逆投影に基づいており、前記第1のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存しないパラメータであり、前記第2のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存するパラメータであり、

前記運動推定ステップにおいて、

- 10       前記処理対象フレームが撮影された撮影時点における前記状態変数の予測値を状態遷移モデルに基づいて算出し、

前記初期投影座標と、前記状態変数の予測値に含まれている前記第1のパラメータと前記第2のパラメータとを、前記モデル座標演算式に適用することによって、前記撮影時点におけるモデル座標の推定値を算出し、

- 15       前記状態変数の予測値に含まれている前記第3のパラメータと前記モデル座標の推定値とを前記運動演算式に適用することによって、前記撮影時点における前記所定の特徴点の座標の推定値を算出し、

前記所定の特徴点の座標の推定値を、前記単眼カメラの観測モデルに基づく観測関数に適用することによって、前記所定の特徴点の観測ベクトルの推定値を算出し、

- 20       前記処理対象フレームから前記所定の特徴点の投影座標を観測ベクトルとして抽出し、

抽出した前記観測ベクトルと前記観測ベクトルの推定値とを用いて前記状態変数の予測値をフィルタリングすることによって、前記撮像時点における前記状態  
25       変数の推定値を算出する、ように前記コンピュータを動作させる、画像処理プログラム。

22. 単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を推定するようにコンピュータを動作させる画像処理プログラムが記録されたコンピュータ読取可能な記録媒体であって、コンピュータに、

第1のパラメータと第2のパラメータと初期投影座標とに基づいて前記所定の特徴点のモデル座標を算出するためのモデル座標演算式における該初期投影座標を前記動画像の初期フレームから抽出する三次元モデル初期化ステップと、

前記初期フレームとは異なる前記動画像の処理対象フレームが撮影された撮影時点での前記所定の特徴点の座標を前記モデル座標から算出するための運動演算式における第3のパラメータと、前記第1のパラメータと、前記第2のパラメータとを含む状態変数の推定値を算出し、該状態変数の推定値に含まれる第2のパラメータに基づいて前記所定の特徴点の運動に関する出力値を出力する運動推定ステップと、  
を実行させ、

前記モデル座標演算式は前記単眼カメラの逆投影に基づいており、前記第1のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存しないパラメータであり、前記第2のパラメータは前記所定の特徴点を含む部分のローカルな運動に依存するパラメータであり、

前記運動推定ステップにおいて、

前記処理対象フレームが撮影された撮影時点における前記状態変数の予測値を状態遷移モデルに基づいて算出し、

前記初期投影座標と、前記状態変数の予測値に含まれている前記第1のパラメータと前記第2のパラメータとを、前記モデル座標演算式に適用することによって、前記撮影時点におけるモデル座標の推定値を算出し、

前記状態変数の予測値に含まれている前記第3のパラメータと前記モデル座標の推定値とを前記運動演算式に適用することによって、前記撮影時点における前記所定の特徴点の座標の推定値を算出し、

前記所定の特徴点の座標の推定値を、前記単眼カメラの観測モデルに基づく観測関数に適用することによって、前記所定の特徴点の観測ベクトルの推定値を算出し、

5 前記処理対象フレームから前記所定の特徴点の投影座標を観測ベクトルとして抽出し、

抽出した前記観測ベクトルと前記観測ベクトルの推定値とを用いて前記状態変数の予測値をフィルタリングすることによって、前記撮像時点における前記状態変数の推定値を算出する、ように前記コンピュータを動作させる、ように前記コンピュータを動作させる画像処理プログラムが記録された記録媒体。

10 23. 単眼カメラで顔を撮像し、撮像した動画像から視線を求める画像処理装置であって、

顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、前記静的パラメータ及び前記動的パラメータを推定することによって視線を求めることを特徴とする画像処理装置。

15 24. 前記静的パラメータは、カメラ座標系における瞳の奥行きであることを特徴とする請求項23に記載する画像処理装置。

25. 前記動的パラメータは、眼球の回転パラメータであることを特徴とする請求項23又は請求項24に記載する画像処理装置。

20 26. 前記眼球の回転パラメータは、眼球の座標系での2つの座標軸に対して回転可能である2自由度であることを特徴とする請求項25に記載する画像処理装置。

27. 単眼カメラで三次元物体を撮像し、撮像した動画像から三次元物体の運動を求める画像処理装置であって、

25 画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータで定義し、前記rigidパラメータ及び前記non-rigidパラメータを推定することによって三次元物体の運動を求めることを特徴とする



画像処理装置。

28. 前記rigidパラメータは、モデル座標系における三次元物体の特徴点の奥行きであることを特徴とする請求項27に記載する画像処理装置。

29. 前記non-rigidパラメータは、モデル座標系における三次元物体の特徴点の変化量であることを特徴とする請求項27又は請求項28に記載する画像処理装置。

30. 単眼カメラで顔を撮像し、撮像した動画像から視線を求める画像処理方法であって、

顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、前記静的パラメータ及び前記動的パラメータを推定することによって視線を求めることを特徴とする画像処理方法。

31. 前記静的パラメータは、カメラ座標系における瞳の奥行きであることを特徴とする請求項30に記載する画像処理方法。

32. 前記動的パラメータは、眼球の回転パラメータであることを特徴とする請求項30又は請求項31に記載する画像処理方法。

33. 前記眼球の回転パラメータは、眼球の座標系での2つの座標軸に対して回転可能である2自由度であることを特徴とする請求項32に記載する画像処理方法。

34. 単眼カメラで三次元物体を撮像し、撮像した動画像から三次元物体の運動を求める画像処理方法であって、

画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータで定義し、前記rigidパラメータ及び前記non-rigidパラメータを推定することによって三次元物体の運動を求めることを特徴とする画像処理方法。

35. 前記rigidパラメータは、モデル座標系における三次元物体の特徴点の奥行きであることを特徴とする請求項34に記載する画像処理方法。

36. 前記non-rigidパラメータは、モデル座標系における三次元物体の特徴点の変化量であることを特徴とする請求項34又は請求項35に記載する画像処理方法。

37. 単眼カメラによって撮像された顔の動画像から視線を求めるようにコンピュータを動作させる画像処理プログラムであって、

顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、前記静的パラメータ及び前記動的パラメータを推定することによって視線を求めるように該コンピュータを動作させる、画像処理プログラム。

38. 単眼カメラによって撮像された三次元物体の動画像から該三次元物体の運動を求めるように、コンピュータを動作させる画像処理プログラムであって、

画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータで定義し、前記rigidパラメータ及び前記non-rigidパラメータを推定することによって三次元物体の運動を求めるようにコンピュータを動作させるように、該コンピュータを動作させる、画像処理プログラム。

39. 単眼カメラによって撮像された顔の動画像から視線を求めるようにコンピュータを動作させる画像処理プログラムが記録されたコンピュータ読取可能な記録媒体であって、コンピュータに読み取られることによって、

顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、前記静的パラメータ及び前記動的パラメータを推定することによって視線を求めるように該コンピュータを動作させる画像処理プログラムが記録された記録媒体。

40. 単眼カメラによって撮像された三次元物体の動画像から三次元物体の運動を求めるようにコンピュータを動作させる画像処理プログラムが記録されたコンピュータ読取可能な記録媒体であって、コンピュータに読み取られることによって、

画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータで定義し、前記rigidパラメータ及び前記non-rigid

パラメータを推定することによって三次元物体の運動を求めるようにコンピュータを動作させる画像処理プログラムが記録された記録媒体。

図1

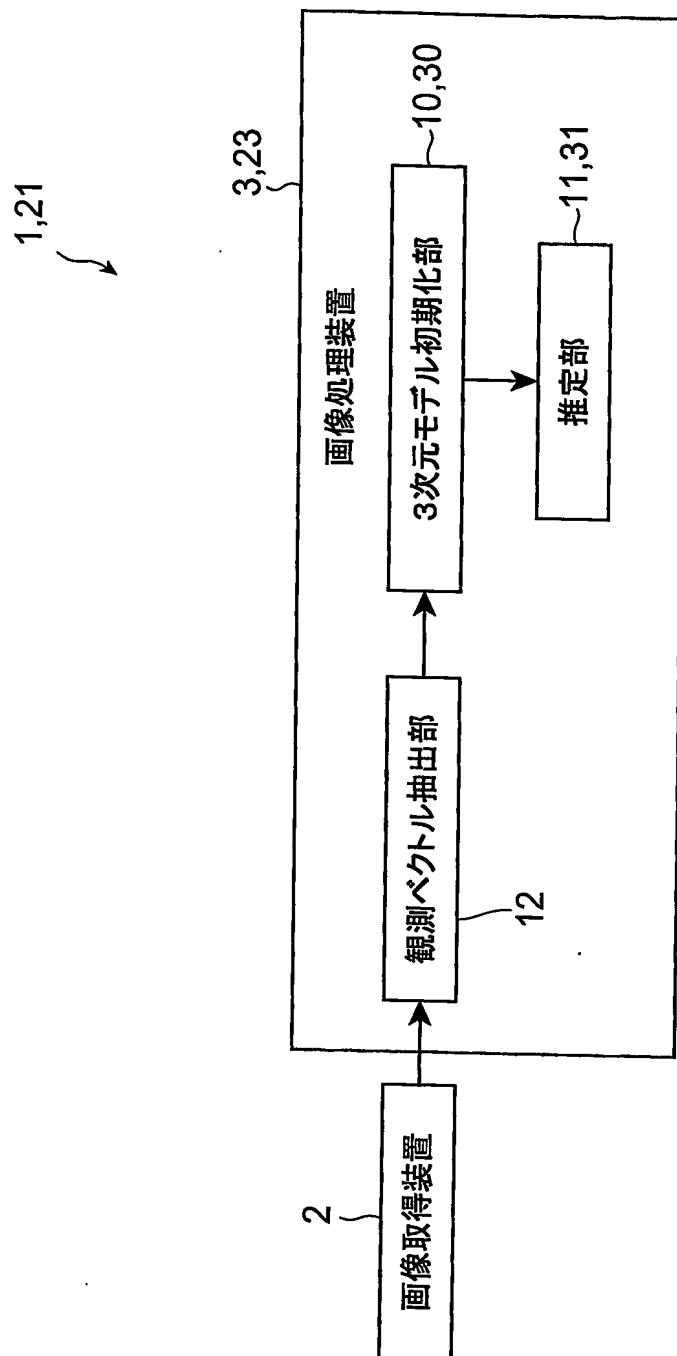


图2

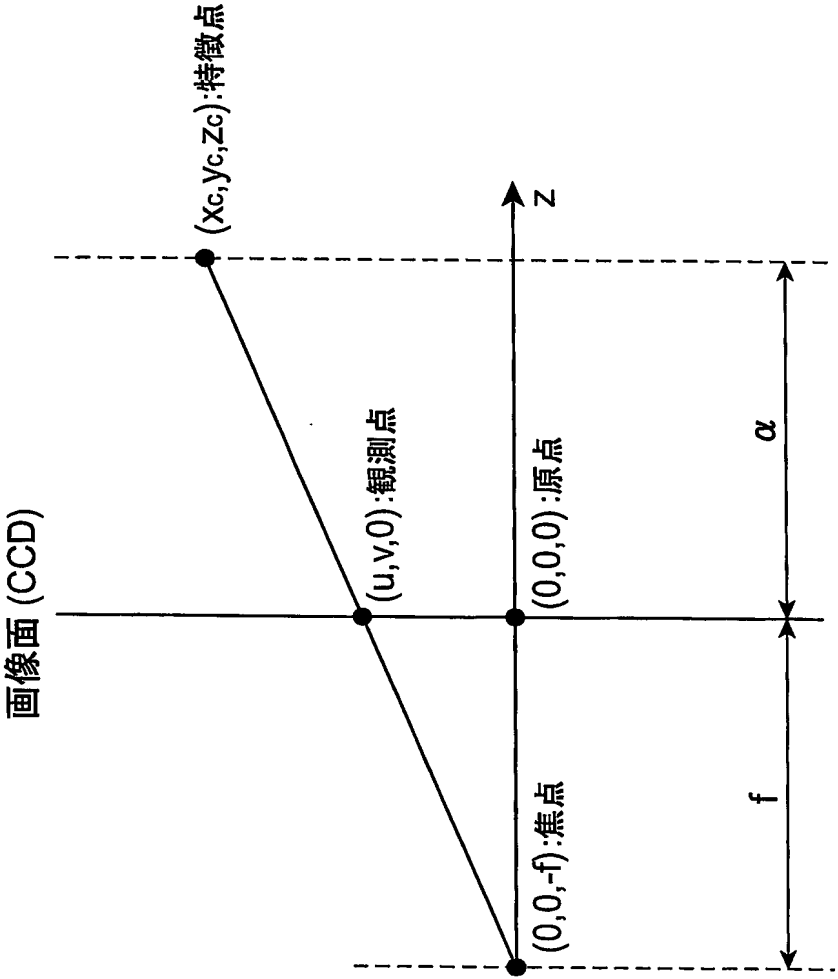


図3

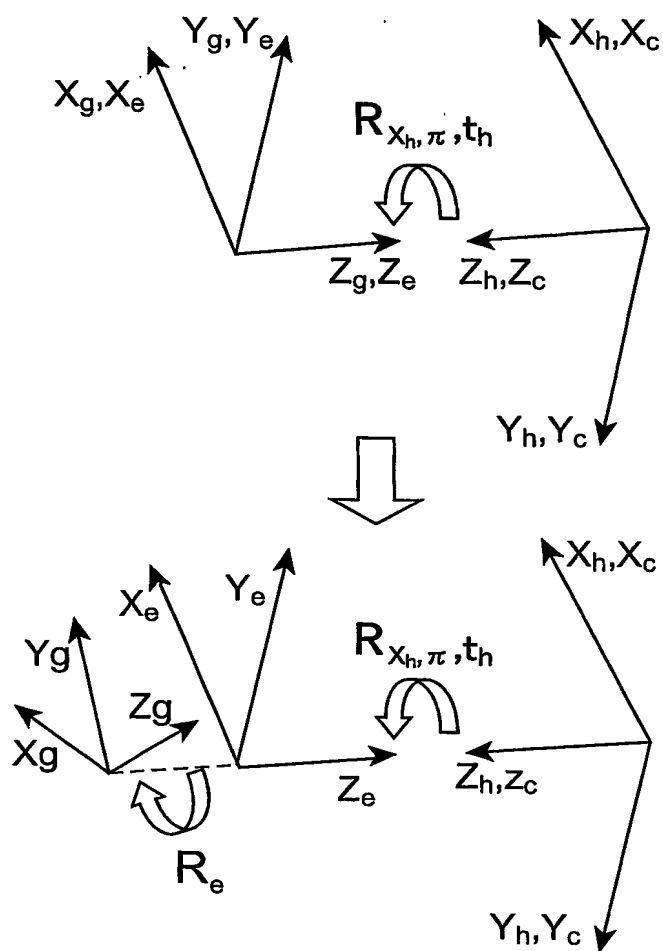


図4

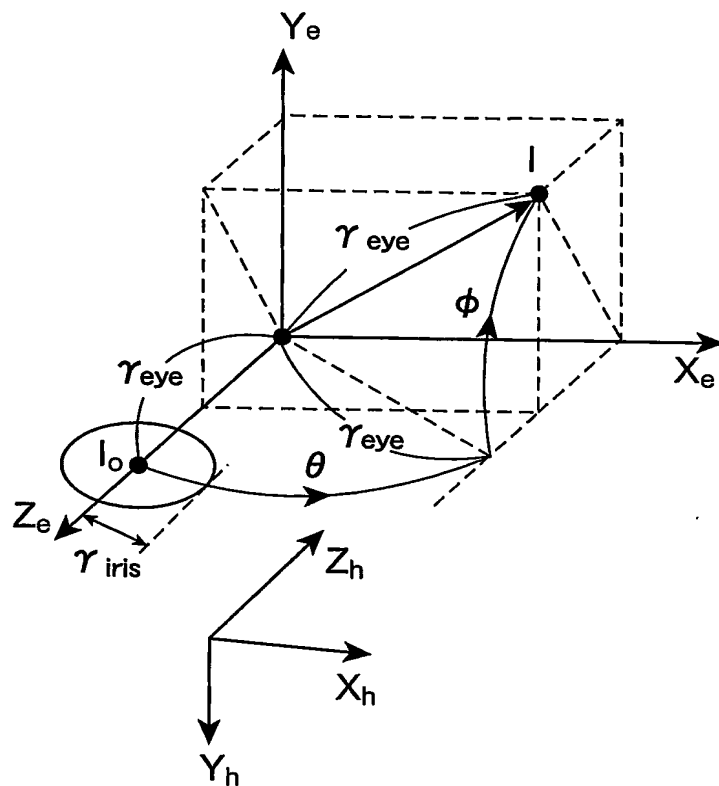


図5

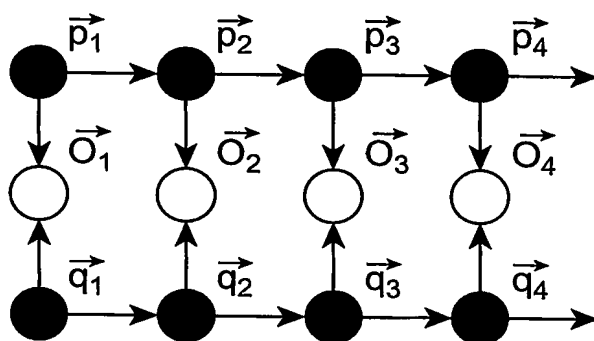




図6

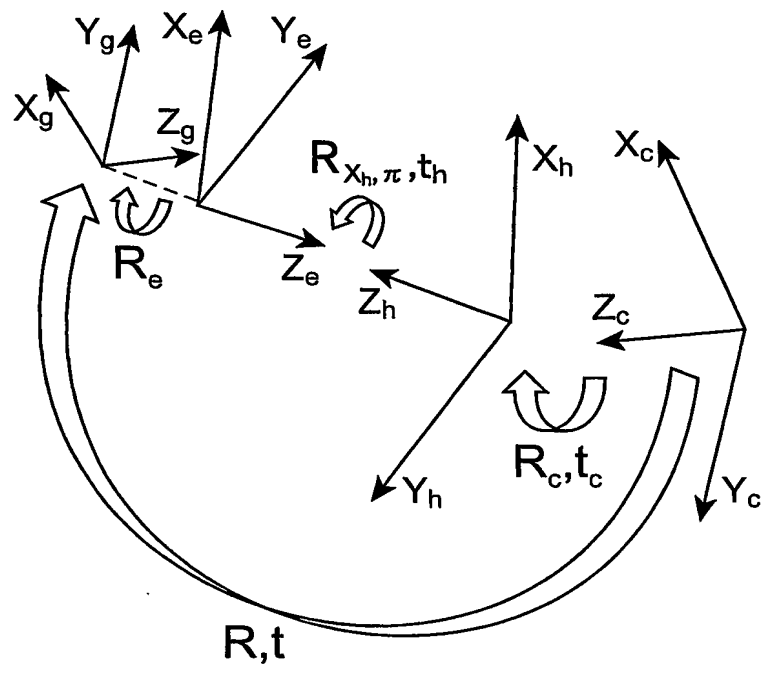


図7

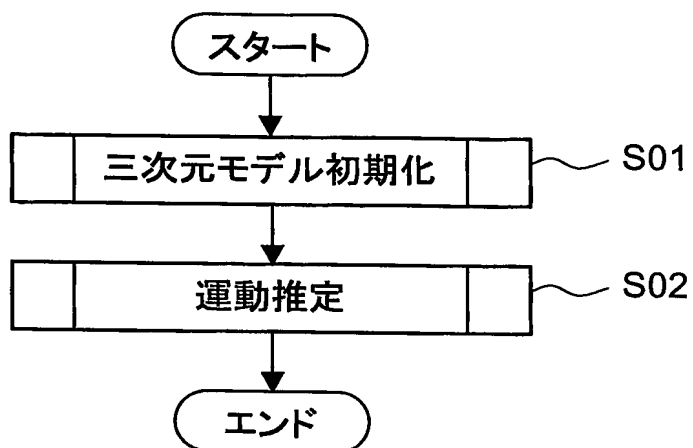


図8

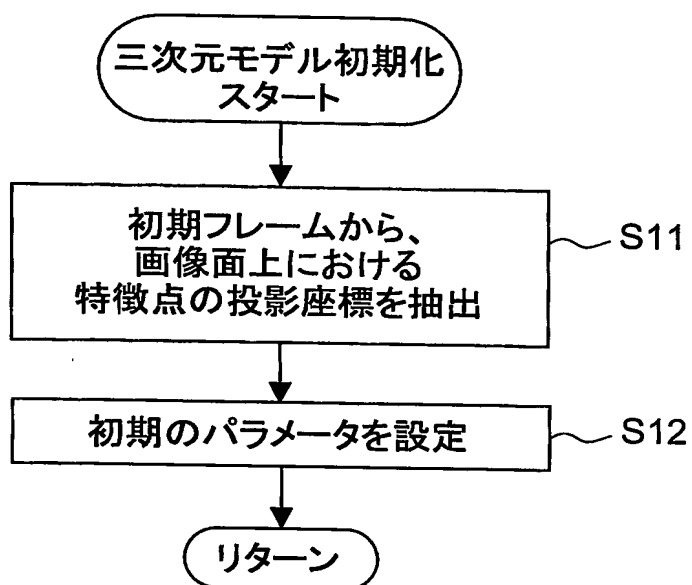


図9

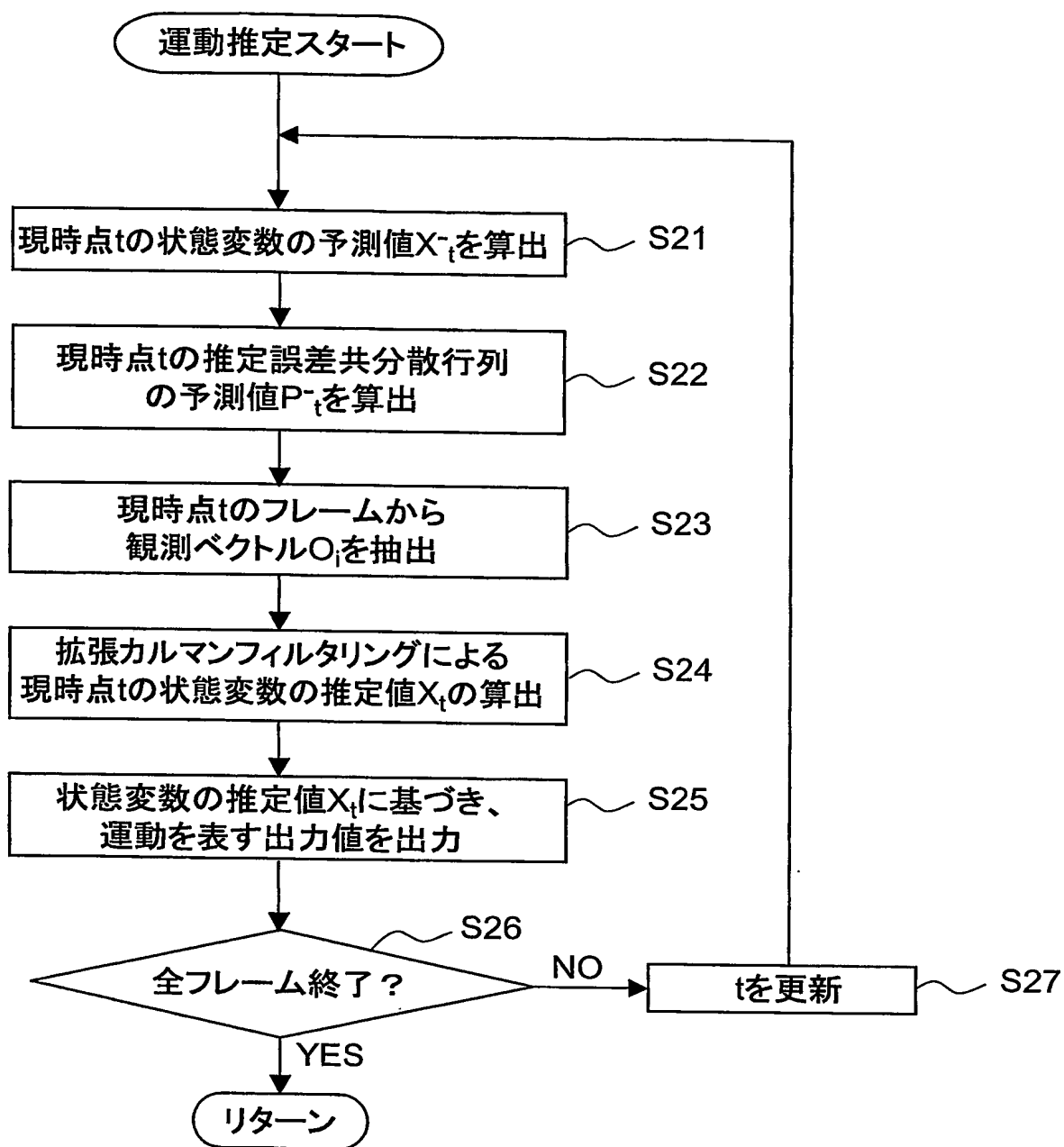


図10

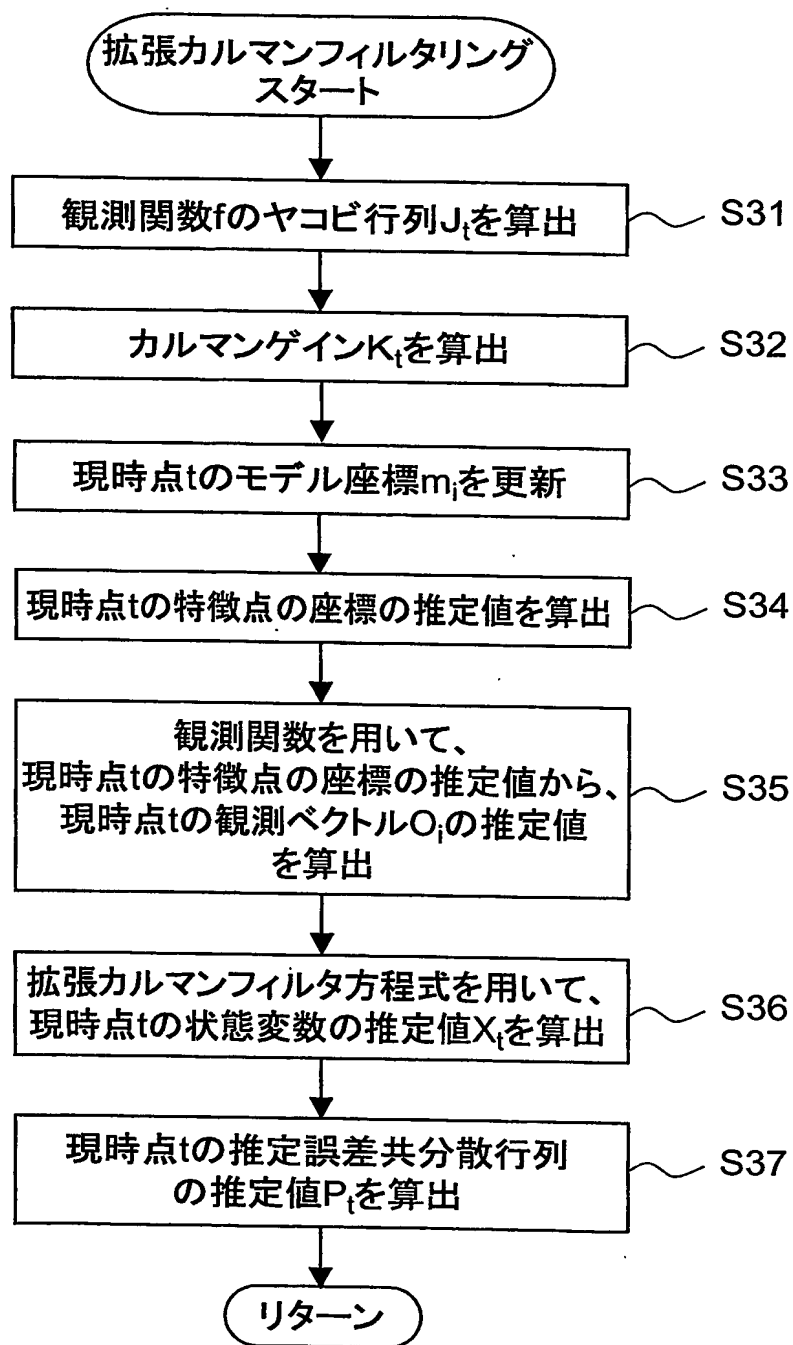


図11

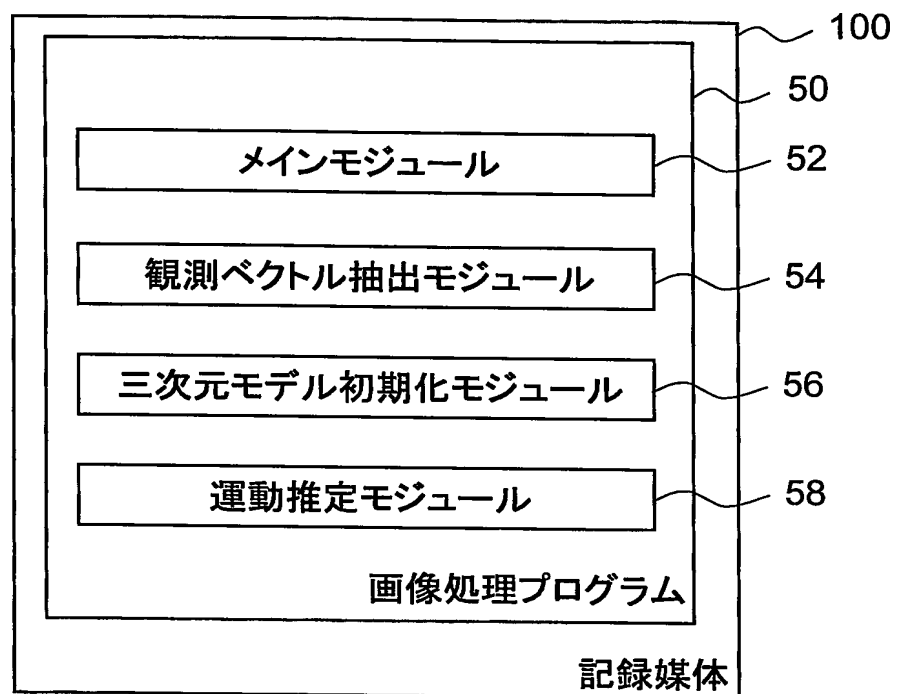


図12

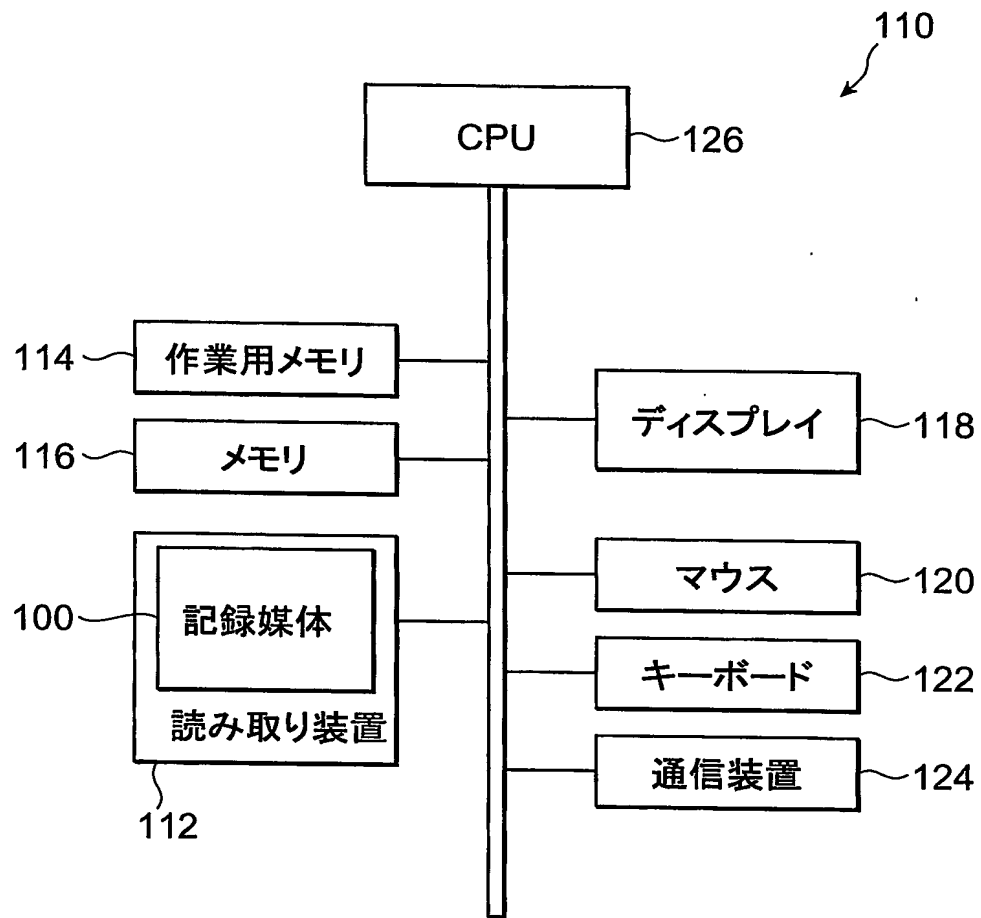


図13

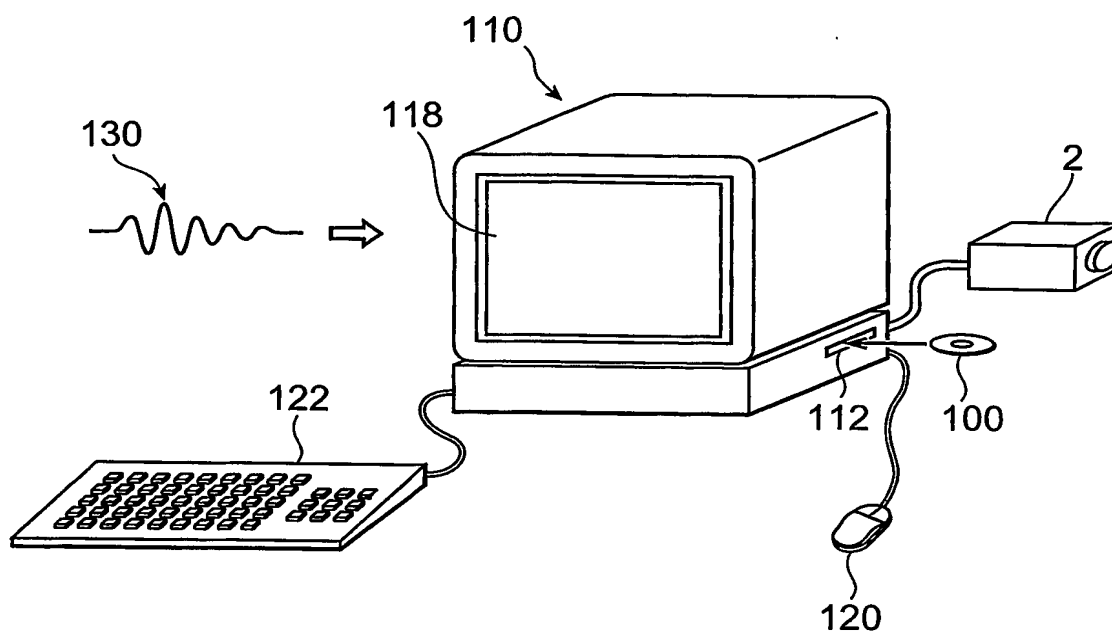




図14A

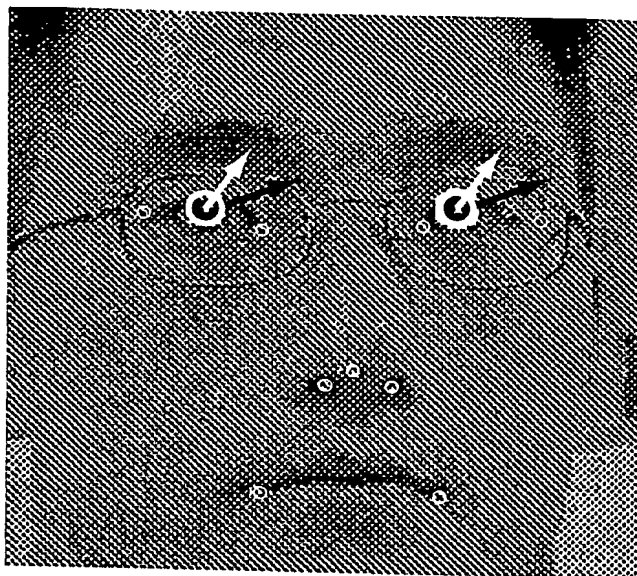


図14B

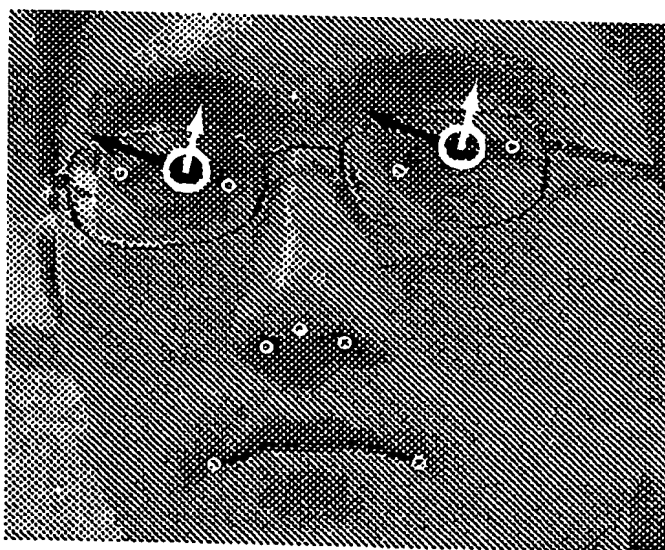


図15A

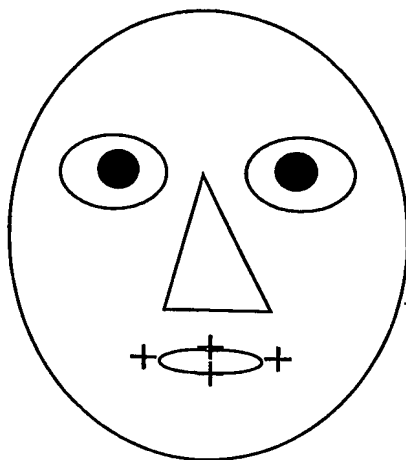


図15B

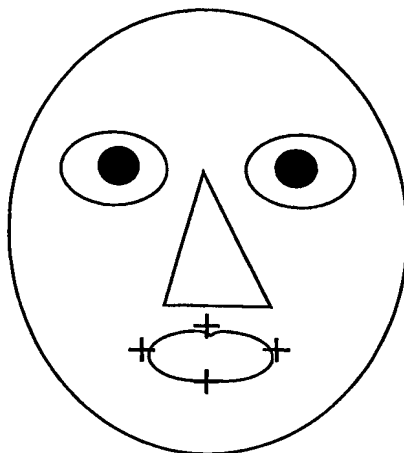


図15C

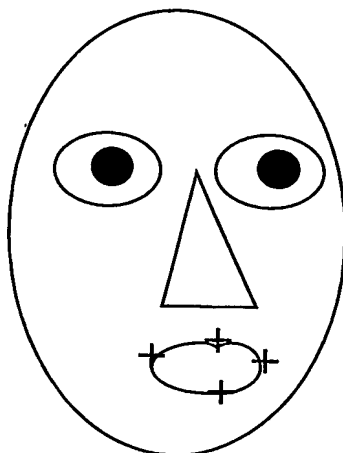
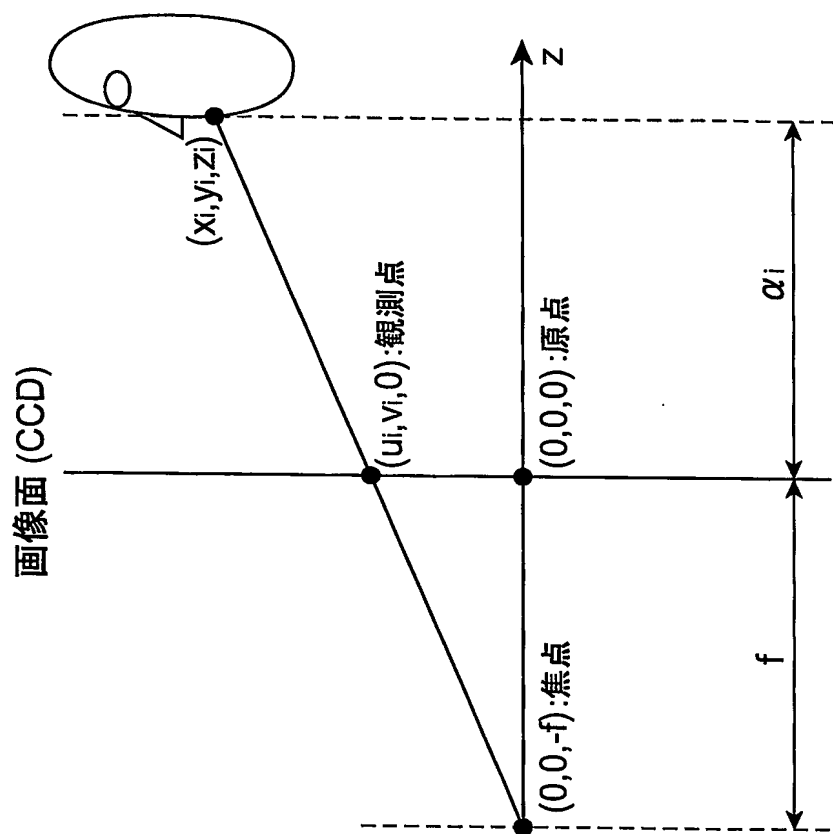


图16



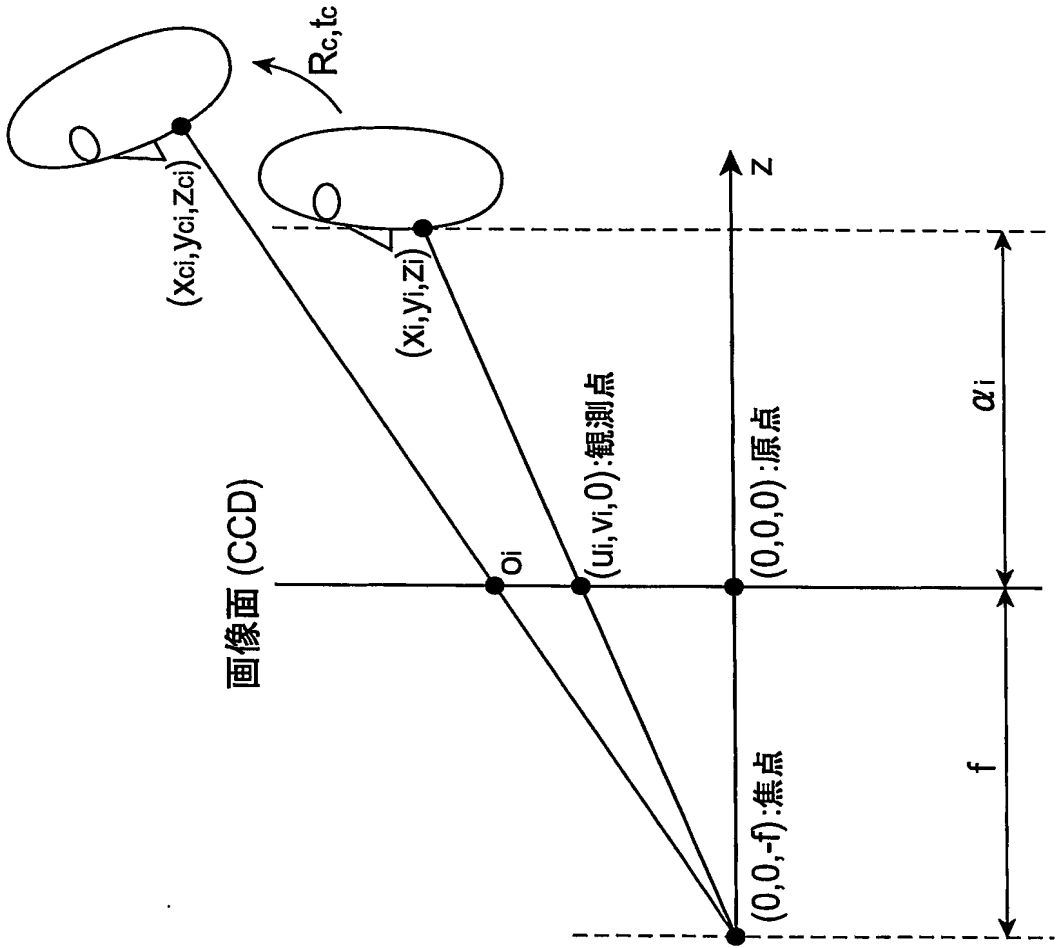


図17

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/010171

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G06T1/00, 7/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G06T1/00, 7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPLUS (JOIS)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 11-281323 A (Isuzu Motors Ltd.), 15 October, 1999 (15.10.99), Par. Nos. [0010] to [0024]; all drawings (Family: none)	23-40
X	JP 11-281324 A (Isuzu Motors Ltd.), 15 October, 1999 (15.10.99), Par. Nos. [0014] to [0039]; all drawings (Family: none)	23-40

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search  
10 August, 2004 (10.08.04)

Date of mailing of the international search report  
31 August, 2004 (31.08.04)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/010171

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Michio MIYAKAWA et al., "Multi Modal Interface Jitsugen ni Muketa Chushiten no Real Time Keisokuho", The Transactions of the Institute of Electronics, Information and Communication Engineers, D-2, The Institute of Electronics, Information and Communication Engineers, 25 December, 2000 (25.12.00), Vol.J83-D-2, No.12, pages 2810 to 2821	1-40
A	JP 08-297019 A (Nissan Motor Co., Ltd.), 12 November, 1996 (12.11.96), Full text; all drawings (Family: none)	1-40

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2004/010171

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

The inventions of claims 1-22 relate to an image processing device including means for extracting observation vectors from the moving picture frames, 3-dimensional model initialization means, and motion estimation means for estimating the motion of a predetermined characteristic point of a 3-dimensional object from the moving picture of the 3-dimensional object imaged by a monocular camera.

(Continued to extra sheet.)

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2004/010171

Continuation of Box No.III of continuation of first sheet(2)

The inventions of claims 23-26, 30-33, 37, 39 relate to an image processing device for imaging a face by a monocular camera, defining a 3-dimensional structure around the pupils on the face image by a static parameter and a dynamic parameter, and estimating the static parameter and the dynamic parameter so as to obtain a visual line from the moving picture imaged.

The inventions of claims 27-29, 34-36, 38, 40 relate to an image processing device for imaging a 3-dimensional object by a monocular camera, defining the 3-dimensional structure of the 3-dimensional object on the image by a rigid parameter and a non-rigid parameter, and estimating the rigid parameter and the non-rigid parameter so as to obtain the 3-dimensional motion of the 3-dimensional object from the imaged moving picture.



A. 発明の属する分野の分類 (国際特許分類 (IPC))  
Int. Cl<sup>7</sup>G06T1/00 7/00

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))  
Int. Cl<sup>7</sup>G06T1/00 7/00

最小限資料以外の資料で調査を行った分野に含まれるもの

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)  
JSTPLUS (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	J P 11-281323 A (いすゞ自動車株式会社) 1999. 10. 15, 第【0010】欄-第【0024】欄, 全図 (ファミリーなし)	23-40
X	J P 11-281324 A (いすゞ自動車株式会社) 1999. 10. 15, 第【0014】欄-第【0039】欄, 全図 (ファミリーなし)	23-40
A	宮川 道夫ほか" マルチモーダルインタフェース実現に向けた注視点のリアルタイム計測法", 電子情報通信学会論文誌D-2, 電子情報通信学会, 2000. 12. 25, Vol. J83-D-2, No. 12, pp. 2810-2821	1-40
A	J P 08-297019 A (日産自動車株式会社) 1996. 11. 12, 全文, 全図 (ファミリーなし)	1-40

☐ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

\* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
「O」口頭による開示、使用、展示等に言及する文献  
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献  
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
「&」同一パテントファミリー文献

国際調査を完了した日  
10. 08. 2004

国際調査報告の発送日  
31. 8. 2004

国際調査機関の名称及びあて先  
日本国特許庁 (ISA/JP)  
郵便番号100-8915  
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)  
脇岡 剛  
5H 9365  
電話番号 03-3581-1101 内線 3531

## 第II欄 請求の範囲の一部の調査ができないときの意見 (第1ページの2の続き)

法第8条第3項 (PCT17条(2)(a)) の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. ☐ 請求の範囲 \_\_\_\_\_ は、この国際調査機関が調査をすることを要しない対象に係るものである。つまり、
2. ☐ 請求の範囲 \_\_\_\_\_ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、
3. ☐ 請求の範囲 \_\_\_\_\_ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

## 第III欄 発明の単一性が欠如しているときの意見 (第1ページの3の続き)

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

請求の範囲1-22に係る発明は、単眼カメラによって撮影された三次元物体の動画像から該三次元物体の所定の特徴点の運動を推定するために、動画像のフレーム各々から観測ベクトル抽出手段と、三次元モデル初期化手段と、運動推定手段とを備えるものである。

請求の範囲23-26, 30-33, 37, 39に係る発明は、単眼カメラで顔を撮像し、撮像した動画像から視線を求めるために、顔画像上の瞳中心の三次元構造を静的パラメータと動的パラメータで定義し、静的パラメータ及び動的パラメータを推定することによって視線を求めるものである。

請求の範囲27-29, 34-36, 38, 40に係る発明は、単眼カメラで三次元物体を撮像し、撮像した動画像から三次元物体の運動を求めるために、画像上の三次元物体の三次元構造をrigidパラメータとnon-rigidパラメータで定義し、rigidパラメータ及びnon-rigidパラメータを推定することによって三次元運動を求めるものである。

1. ☐ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
2. ☒ 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったため、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったため、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

追加調査手数料の異議の申立てに関する注意

- ☐ 追加調査手数料の納付と共に出願人から異議申立てがあった。
- ☐ 追加調査手数料の納付と共に出願人から異議申立てがなかった。

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ ~~FADED~~ TEXT OR DRAWING
- ☐ ~~BLURRED~~ OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**